广播电视新闻系列教材

精确新闻学

JING QUE XIN
WEN XUE

肖明 丁迈 著



中国广播电视出版社
ZHONG GUO GUANG BO DIAN SHI CHU BAN SHE

广播电视新闻系列教材

精确新闻学

PRECISION JOURNALISM

肖明　丁迈　著

中国广播电视出版社

图书在版编目（CIP）数据

精确新闻学/肖明，丁迈著．–北京：中国广播电视出版社，
2002.1

ISBN 7–5043–3854–0

Ⅰ.精… Ⅱ.①肖… ②丁… Ⅲ.新闻学 Ⅳ.G210

中国版本图书馆 CIP 数据核字（2001）第 092644 号

精确新闻学

作　　者：	肖　明　丁　迈
责任编辑：	贺　明
封面设计：	李燕平
责任校对：	谭　霞
监　　印：	戴存善
出版发行：	中国广播电视出版社
电　　话：	86093580　86093583
社　　址：	北京复外大街 2 号（邮政编码　100866）
经　　销：	全国各地新华书店
印　　刷：	廊坊人民印刷厂
装　　订：	廊坊人民印刷厂
开　　本：	850×1168 毫米　1/32
字　　数：	290（千）字
印　　张：	14.5
版　　次：	2002 年 1 月第 1 版　2002 年 1 月第 1 次印刷
印　　数：	4000 册
书　　号：	ISBN 7-5043-3854-0/G·1507
定　　价：	25.00 元

内容摘要

精确新闻是指记者在采访新闻时，运用调查、实验和内容分析等社会科学研究方法，来收集资料、查证事实，从而报道新闻。本书共分为十二章，系统地介绍了与精确新闻学有关的理论和方法。具体如下：

第一章　精确新闻的发展历史以及现状；第二章　精确新闻学中涉及的概念；第三章　抽样方法；第四章　内容分析方法；第五章　实验方法；第六章　调查方法；第七章　电脑辅助的电话调查 CATI；第八章　调查问卷的设计；第九章　数据的整理；第十章　SPSS for Windows 基本操作；第十一章　常用统计分析方法；第十二章　精确新闻的写作。

本书的特点是内容丰富，概念清晰，并列举了大量的实例，强调实际应用。

本书可供各类高等院校的新闻专业的学生作为教材或参考书使用，也可以作为新闻从业人员的参考书或培训教材。



序言一

有学者指出，工业社会的特征是细分，知识经济时代的特征是整合。近半个世纪以来，随着社会进步、技术发展，自然科学与社会科学不断渗透融合，相继出现了许多新的边缘学科。精确新闻就是借用社会科学的研究方法，将调查结果借助计算机等现代技术进行分析和处理，运用“数字语言”对新闻事实进行定量的描述、分析，力求使新闻报道更加精确、客观、可信。80年代德国新闻学者麦库姆斯把自己的专著定名为《精确新闻学——一种新出现的新闻报道理论和技巧》，该书的副题“一种新出现的新闻报道理论和技巧”从一个侧面对“精确新闻学”的社会背景和技术背景进行了阐释，也为我们更加全面地理解精确新闻提供了一个参照系。

“它山之石可以攻玉”，近年来内地新闻传播学者从学科交叉兼容的视角对精确新闻学理论与实务从不同侧面进行了研究，不少作品颇有见地。近年来国内主流传媒对精确新闻报道亦颇为重视，出现了一批有较大影响的专栏和“品牌”。这些成果从理论到实践为本书作者提供了可借鉴的思路和启示。

肖明和丁迈是北京广播学院新闻传播学院两位年轻的教师，多年从事统计学、受众研究、市场调查与统计分析等课程



的教学，积累了丰富的教学经验，并取得了良好的效果。她们突破传统学院式研究的局限，重视理论实际的结合，多次承担或主持大型的社会调查统计课题，有的课题在业内具有较大的影响。

本书两位作者的学科背景都属于理工科，接触新闻传播学的时间还不是很长。由于知识背景的缺陷，书中难免存在需要改进之处。用数字解读新闻事实在一定程度上还是表面化的读解，对于业者而言，更重要的是要让受众学会读解数字背后的深刻内涵。在这一点上，两位作者还需进行更为深刻的研究和实践，以从理论和实践两方面对精确新闻学做进一步的深化和拓展。

由于本人知识结构的局限，对新的研究方法在新闻传播学的运用似不具发言权，以上文字权当老教师对“青出于蓝而胜于蓝”的期待。

曹 璐

2001 年 10 月 1 日



序言二

改革开放以来，我国的媒体在精确新闻报道方面有了长足的进步，并受到了广泛的重视，不少媒体设立了诸如“新闻调查”、“市场调查”、“民意”、“精确新闻”、“每周调查”等报纸专版或广播、电视固定栏目。记者通过科学的研究方法和手段，收集资料、查证事实，从而更加准确、客观地报道新闻。精确新闻报道逐渐受到了理论界和教育界的认可和受众的欢迎。

然而，如何将起源于西方的精确新闻学很好地应用于我国的新闻实践，并不是一件容易的事情。几年前，我曾让学习我所讲授的“传播研究方法”的研究生，去查阅和分析国内主要报刊上某个时期内所刊登的几百篇与精确新闻有关的报道，结果发现：真正符合精确新闻报道要求的文章简直是凤毛麟角。虽然这可能与个别记者自身的素养有关，但是我认为主要的原因，还是由于我国新闻人才的教育培养体系中，缺少对精确新闻学理论方法的系统学习和训练。我期盼着能有一本内容丰富、深入浅出、强调实际应用的精确新闻学教材早日问世。

现在看到了这本《精确新闻学》的出版，本书的作者是与我共事近10年的两位新秀肖明和丁迈。这些年来，肖明和丁

迈一直从事与社会研究方法、受众调查和统计学有关的教学、科研和实践。她们不但理论功底较为扎实，而且有着相当丰富的实践经验；特别让我感到欣慰的是，她们有执著的事业心和高度的责任感。这本书是她们的新作，为了奉献给读者更有价值的东西，她们查阅和参考了大量的国内外资料和文献，并结合了自己的实践经验和研究成果，所以本书的内容相当丰富。

这部教材肯定会帮助有志于精确新闻报道的读者，在较短的时间内掌握与精确新闻有关的基本概念，并学会科学地获取精确新闻资料的主要方法，以及认识客观地报道精确新闻的基本原则。这部教材对于目前从事新闻工作或是对精确新闻有兴趣的读者，也一定会有所裨益。相信这本新书的出版定会给我国新闻学的教育以及新闻学的建设带来一些新的思考。

柯惠新

2001 年 9 月 15 日





作者简历

肖明，北京广播学院新闻传播学院调查统计研究所教师，曾在中国人民大学和香港城市大学学习、工作。

近十年来，一直从事数据处理、市场调查、受众研究等方面的教学与科研工作，主讲过《社会统计学》、《受众研究方法》、《市场调查与分析》、《精确新闻学》等课程，在《现代传播》、《数理统计与管理》、《统计信息与论坛》等学术杂志上发表近20篇论文。

研究方向为：精确新闻学、受众研究、市场调查、统计分析以及抽样设计。

丁迈，北京广播学院新闻传播学院教师，毕业于北京师范大学物理系，在北京广播学院获得通信与信息系统专业硕士学位。1995年开始从事市场调查、统计分析、受众研究等方面的教学与研究工作，在校期间讲授了《普通物理》、《线性代数》、《社会统计学》、《传播研究与统计分析方法》、《市场调查与分析》等课程，曾在《有线电视技术》、《数理统计与管理》等行业杂志、报刊、会议上发表了数篇论文，并与肖明等合作翻译了《传播研究方法：策略与资料来源》一书。

主要研究方向为：精确新闻学、传播研究、市场调查、统计分析。



目　录

序言一 ………………………………………………………… (1)
序言二 ………………………………………………………… (3)

第一章　什么是精确新闻？ ……………………………… (1)
第一节　精确新闻的定义…………………………………… (2)
第二节　精确新闻学的兴起和发展………………………… (9)
第三节　美国主要媒体从事民意调查的情况 ………… (22)
第四节　精确新闻在中国的发展 ……………………… (25)
第五节　报道精确新闻的程序 ………………………… (28)
应用实例 ……………………………………………………… (30)

第二章　精确新闻学的基本概念 ……………………… (35)
第一节　研究的类型 ……………………………………… (37)
第二节　概念：定义及其操作化 ……………………… (48)
第三节　变量的类型及转换 …………………………… (55)
第四节　量表 ……………………………………………… (58)
第五节　测量的信度和效度 …………………………… (63)

第三章　抽样 ……………………………………………… (69)
第一节　抽样调查的基本原理 …………………………… (72)
第二节　抽样调查的基本概念 …………………………… (80)
第三节　概率抽样的几种基本形式 ……………………… (93)
第四节　非概率抽样的几种基本形式…………………… (103)

第四章　内容分析法 ………………………………… (111)
第一节　什么是内容分析？ …………………………… (111)
第二节　内容分析的功能………………………………… (116)
第三节　内容分析的研究程序…………………………… (119)
第四节　信度与效度……………………………………… (127)
应用实例…………………………………………………… (130)

第五章　实验法……………………………………………… (142)
第一节　实验法的基本概念……………………………… (144)
第二节　实验设计………………………………………… (151)
第三节　内部效度和外部效度…………………………… (163)
应用实例…………………………………………………… (172)

第六章　调查法……………………………………………… (178)
第一节　面访调查………………………………………… (179)
第二节　邮寄问卷调查…………………………………… (183)
第三节　电话调查………………………………………… (190)
第四节　三种调查访问方式比较………………………… (192)
第五节　大众媒介在发表调查结果时应注意的问题
……………………………………………………… (194)



第七章 电脑辅助的电话调查 CATI …………………… (202)

第一节 CATI 系统的访问过程 ……………………… (203)

第二节 电话调查的抽样 …………………………… (215)

应用实例 ……………………………………………… (222)

第八章 调查问卷的设计 ………………………… (231)

第一节 问卷的作用与类型 ………………………… (231)

第二节 问卷设计的基本步骤 ……………………… (237)

第三节 问卷中常用题型和量表的设计 …………… (253)

第四节 电话调查问卷的设计 ……………………… (264)

第九章 调查数据的整理 ………………………… (273)

第一节 调查资料的接收、检查和校订 …………… (274)

第二节 调查数据的编码和录入 …………………… (278)

第三节 数据审核 …………………………………… (287)

第四节 对数据的统计预处理 ……………………… (294)

第十章 SPSS for Windows 基本操作 …………… (303)

第一节 SPSS 的基本特点和功能介绍 …………… (303)

第二节 变量及其相关操作 ………………………… (320)

第三节 数据输入和数据文件编辑 ………………… (327)

第十一章 常用统计分析方法 …………………… (344)

第一节 统计概论 …………………………………… (344)

第二节 描述性统计 ………………………………… (346)

第三节 参数估计和假设检验 ……………………… (364)



第四节　常用双变量分析方法 …………………………… (385)
第五节　常用统计图表 …………………………………… (415)

第十二章　精确新闻报道的写作技巧 …………… (424)
第一节　精确新闻倒金字塔式写作模式 ……………… (424)
第二节　精确新闻钻石式写作模式 …………………… (434)
第三节　精确新闻基本形态及写作中应注意的问题
…………………………………………………… (439)

参考文献 ……………………………………………… (444)
后记 …………………………………………………… (448)



第一章 什么是精确新闻?

精确新闻是最近几年在我国媒体上出现的一种报道方式，首先，我们来看一则《北京晚报》2001年8月22日第五版上的精确新闻报道。

第四次“中国信贷消费调查”今天发布结果
大部分居民不敢透支

[**本报讯**]（记者辛宏）由中国经济景气监测中心与万事达卡国际组织就居民信贷消费有关问题进行的第四次调查，于今天公布结果。此次调查显示，我国城市居民对信贷消费的认知度有所提高，绝大部分人对个人经济预期信心不足，不敢使用信用卡的透支功能。

此次对北京、上海、郑州、武汉、广州、成都和西安7个城市2100名居民进行的问卷调查显示，居民对信贷消费概念的认知度又有了进一步提高，认可信贷消费方式的被访者比例比今年一季度上升了2.7个百分点。

有关银行卡的专项调查发现，大部分（73.6%）持有信用卡的被访者从未使用过透支功能，经常使用的被访者比例仅为4.2%。高收入者比低收入者更喜欢使用信用卡的透支功能。

(《北京晚报》2001年8月22日第五版)

在同一天的《北京晚报》上，第六版还有另一篇精确新闻报道《向往身体健康心情舒畅，北京市民“幸福观”发生变化》。

许多精确新闻报道都是以政治作为主题的。

“台北市长马英九上任将满3个月，市府研考会上星期对市府团队上任3个月的施政满意度进行民意调查，受访者对马英九整体表现感到满意的有六成五，对市府团队整体表现感到满意的只有四成多。

对马英九上任以来的治安，有过半数的受访者认为治安有进步，由于近来有不少民代要求市府应开放合法的电动玩具，建设局也将以3个月时间评估电动玩具要不要开放；大部分的受访者对是否应开放电玩，都持反对意见。”(摘自台湾《联合报》1999年3月22日第四版〈政治版〉)

不仅印刷媒体报道精确新闻，在广播电视媒体上，也经常出现精确新闻报道。

在本章我们将介绍精确新闻的概念、发展和现状。

第一节　精确新闻的定义

一、精确新闻学的定义

精确新闻学（Precision Journalism)，也称精确新闻报道、精确报道，是由美国学者、新闻记者菲利普·迈耶（Philip Meyer）最早提出来的，是指记者在采访新闻时，运用调查、实验和内容分析等社会科学研究方法，来收集资料、查证事实，从而报道新闻。



1981 年德国新闻学者麦库姆斯，在《精确新闻学：一种新出现的新闻报道理论和技巧》(Precision Journalism: An Emerging Theory and Technique of News Reporting) 一书中，给精确新闻下了这样的一个定义：

"所谓精确新闻学，就是要求记者用科学的社会观察的方法去采集和报道新闻，这些方法包括：民意调查、内容分析、亲自观察和现场实验等等。"

1987 年 David Pearce Demers 与 Suzanne Nichols 在《精确新闻学——实务指南》(precision journalism-A practical Guide) 一书中认为："精确新闻学是一种探寻世界的方法，只不过是采用定量的社会科学研究方法来获得新闻。这里的定量是指把事件、人物、行为或态度用数字来表示，然后进行分析。精确新闻中最常用的两种方法是内容分析法和调查研究法，内容分析法是以有关的文件或记录为研究对象，而调查研究法是以人作为研究对象。另外，实地实验的方法也曾被用到。"[①]

在《新闻学、传播学新名词词典》中，对精确新闻的解释如下：

"精确新闻学是美国新闻学的一个支派，代表作为 1971 年出版的《精确新闻学》一书。该书作者迈耶 (Philip Meyer) 曾长期从事新闻工作。他认为，传统的搜集和加工新闻信息的方式已经过时，因为它普遍地对事实描写得不准确，评价似是而非，一味追求瞬间即息的、耸人听闻的情节，很少从社会历史的角度和政治角度去分析材料，发现具有重大意义的社会现象。他主张新闻记者要广泛使用社会科学的方法和手段，以保

---

① David Pearce Demers 与 Suzanne Nichols《precision journalism-A practical Guide》第 10 页。

证新闻的准确和客观。这些方法有：运用计算机进行选择性调查、内容分析、中间测量、数量分析等。《精确新闻学》出版后，这一学派教科书和各种参考书在美国流行一时。

精确新闻学这一概念虽然在 1971 年以后出现，但在新闻报道中突出数学和统计数字的作用，多用社会调查来说明问题，却在百余年前就出现了。1810 年，《北卡罗来纳州明星报》的编辑亨得森（Thomas Henderson Jr.）和琼斯（Calvin Jones）对该州几个不同地区民众进行一项问卷调查，询问有关农产品和社会福利方面的意见。这一事件被认为是世界新闻史上第一次民意调查，也是精确新闻学的最早实践。其后，1824 年由《哈里斯堡宾州人报》首创的预测总统选举调查，1851 年密西西比州的《纳奇兹快讯报》举办的对 800 选民大规模调查和 1932 年盖洛普发明科学抽样调查法，为精确新闻学的形成提供了有益资料。"[①]

从实践的角度来看，民意调查是新闻界最经常使用的社会科学研究方法，所以也有学者认为"精确新闻报道是一种运用民意测验方法进行新闻素材的收集、整理、分析和报道的一种新闻报道样式。与传统的新闻报道手法不同，它是一种对于社会情况和公众意见的严谨、科学的系统采集，它给予受众的不是个别的、片断的、局部的和割裂的现象描述及基于这种描述的分析，而是一种客观的、全面的、结构化的现象和意见的描述和分析。因此，精确新闻学能够提供社会的标准意见，起到决策参考、社会示范和社会沟通的作用，使新闻报道更加客观精确，民意表达更加畅达明晰等"[②]。

---

① 陶涵 主编《新闻学、传播学新名词词典》，经济日报出版社，1997 年，第 86 - 87 页。

② 喻国明 著《解构民意》，华夏出版社，2001 年，第 38 页。



用社会科学的方法来收集数据、报道新闻，还有一些别的叫法，如电脑辅助的新闻学（computer-assisted journalism）、科学新闻学（scientific journalism）、定量新闻学（quantitative journalism）等，但都不如精确新闻学准确[1]。

二、精确新闻的作用

精确新闻的作用表现在3个方面：

1. 判断消息的真伪，摆脱消息来源的控制，使新闻报道客观准确

传统的新闻报道方式中，记者在采访新闻时，完全依赖于消息来源提供的材料。在现今社会，各种团体、个人都可能出于有利己的目的，有意、无意地向媒体透露不实或者是片面的信息。多数时候，记者面对消息来源所提供的信息，无法辨别真假，也难以了解有关的背景，所以，使得记者容易受消息来源的摆布，被人利用，甚至会发生错误。

一个经典的例子就是美国新闻界1969年报道黑豹党事件时所发生的错误。1969年12月，芝加哥便衣警察在搜查非法枪械时，枪杀了两名黑豹党成员。几天后，黑豹党发言人向新闻界宣布，过去两年内，黑豹党已有28名成员被警方杀害。美国的三大广播网以及纽约时报、华盛顿邮报等媒体均报道了黑豹党对警方的指控。

但是著名的媒介批评家爱泼斯坦（Edward J.Epstein）分析了法院和警方的记录，发现黑豹党所说的28人中，有的是因为抢劫被店主枪杀，有的是被黑豹党成员杀死，有的是因为殴

---

① David Pearce Demers 与 Suzanne Nichols《precision journalism-A practical Guide》第12页。

打太太，被太太杀死。被警方杀害的只有两位。

另一个例子是1976年美国纽约警方打算成立一支打击犯罪的特勤队伍，希望得到广大市民的支持。于是警方在遇到老年人受到伤害的案件时，立即向新闻界公布。一时间，各个报纸和电视台连续不断地报道伤害老年人的案件，使得人人感到危机四伏，警方成立特勤队的打算获得了广大市民的支持，顺利成立，达到了警方的目的。而通过查阅当时警方的记录，就会发现实际上伤害老年人的犯罪事件并没有增多，只是警方对新闻界公布的多了，新闻界的报道也就增多了。

2. 反映普遍的民意，实现舆论监督

在传统的新闻报道中，遇有重大新闻事件时，由于时间、资源等方面的限制，只能采访少数专家学者、社会知名人士、政府官员等，了解他们对新闻事件的看法，而无法反映普通老百姓的意见。

掌握了精确新闻的报道方法后，就可以通过民意调查的方式，迅速获取公众的意见，反映普遍的民意。精英的意见和民众的意见结合起来，可以帮助受众更好地了解新闻。

在2001年中秋节前夕，有记者对南京冠生园月饼“老馅新做”的黑幕进行了曝光，一时间舆论哗然，不仅有食品监督部门、质量检测部门的意见，也有广大的食品厂的看法建议，当然，更少不了月饼的消费者——广大民众的意见。“调查结果显示，有近一半的被调查者对市场上销售的月饼质量、卫生状况表示了不同程度的不放心，但仍有一多半的被调查者表示今年中秋可能还会买月饼。”（《北京青年报》2001年9月19日精确新闻版）。

新闻记者被誉为“无冕之王”，具有监督政府部门运作、评估施政得失的责任。对于有关政策的制定，新闻界可以利用



调查来了解民意。在国外，媒介常常进行有关总统或政府官员的声望或支持率调查，并加以报道，从而加强新闻界对政府部门的监督。

3. 验证现有观点，澄清错误认识

有时社会上所公认的某些看法、观念可能是错误的、片面的,通过科学的方法,可以验证这些观点,从而澄清错误认识。

20世纪60年代开始,在美国不断有种族骚乱发生,开始时是在洛杉矶的Watts地区,后来又扩展到纽瓦克和底特律。当时很多人都认为是那些在社会经济的底层的充满挫折和失败的、毫无希望的黑人参加了骚乱,因为他们没有别的发泄的机会。

这个观点就可以通过调查研究的方式来验证。把参加骚乱的黑人和没有参加的黑人进行比较，如果这个观点是正确的话，那么受教育程度低的人将会倾向于参加骚乱。1967年由《底特律自由报》（Detroit Free Press）的菲利普·迈耶主持的一项调查表明，那些上了大学的黑人与那些在中学就退学的黑人在参与骚乱方面没有差别。具体的调查结果请见表1-1:[1]

表1-1

比　　例	文化程度		
	退　学	高　中	大　学
参加了骚乱	18%	15%	18%
没有参与骚乱	82%	85%	82%
合　　计	100%	100%	100%

① 菲利普·迈耶《New Precision Journalism》，第16页。

从调查结果可知，在文化程度为大学毕业的黑人中，有18% 的人参加了骚乱，这个比例与那些中学就退学的黑人中参加骚乱的比例相同。从而说明了，一个黑人是否参加骚乱与受教育程度的高低是没有关系的。

当时，另一种被公认的观点认为骚乱的根源在于，那些成长于南方的黑人难以被北方的文化同化，在南方奴隶制的压迫下，南方的黑人形成了消极被动的性格，一旦他们离开南方，骚乱就成为他们发泄长期被压抑的情绪的方式。这种说法也容易被检验，这种说法如果是正确的，那么从南方移民过来的黑人比那些在北方土生土长的黑人更容易参加骚乱。

菲利普·迈耶共调查了 437 人，其中在南方长大的有 237 人，在北方长大的有 200 人，见表 1-2：[①]

表 1-2

人数（比例）	你在哪儿长大的？		
	南　方	北　方	合　计
参加了骚乱	19（8%）	51（25%）	70
没有参与骚乱	218（92%）	149（75%）	367
合　计	237（100%）	200（100%）	437

调查结果显示，在南方长大的黑人中，只有 8%参与了骚乱，而在北方土生土长的黑人却有 25%参与了骚乱。调查结果否定了流行的观点。

另一个例子就是 1968 年黑人领袖马丁·路德·金被暗杀，

① 菲利普·迈耶《New Precision Journalism》，第 53 页。



当时舆论都认为，随着马丁·路德·金的去世，他所倡导的非暴力运动也会随他而逝，黑人们会转而拥护倡导暴力的领袖。《迈阿密先驱报》（Miami Herald）在马丁·路德·金被暗杀前刚好完成了一项调查，目的在于了解黑人对这件事情的态度，马丁·路德·金被暗杀后，他们马上找到原来的被访者，再次进行调查，结果发现大家更加反对暴力，拥护马丁·路德·金的非暴力运动的主张更加坚定了。

精确新闻的巨大作用，使得新闻从业人员越来越重视学习精确新闻技能的重要性。

1989年10月在华盛顿的一次会议上，《今日美国》（USA Today）主编 Julia D.Wallace 刚刚学习完几个月的精确新闻课程，她深有感触地说："我猜想在未来要成为一个新闻工作者的底线将会更高，具有良好的写作能力、广泛的消息来源以及旺盛的精力是不够的，记者和编辑都需要掌握一些统计以及软件的知识。虽然掌握这些不是很容易，但学到最后，会发现这是有趣的。"[①]

第二节 精确新闻学的兴起和发展

精确新闻学兴起、发展于美国，民意调查是精确新闻报道的最主要组成部分，所以，精确新闻学的兴起与发展是以民意调查在美国的发展为主线的。

---

① Philip Meyer《New Precision Journalism》，第17页，原文如下：
"I guess the bottom line here is it is going to take more to be a journalist in the future. It's not enough to have writing ability, good sources, and lots of energy. Reporters and editors are going to need to know something about statistics and software. It's not easy. But in the end, it sure is fun."

一、早期的精确新闻报道研究

早在1810年，美国就有报纸尝试着进行精确新闻报道。《北卡罗来纳州明星报》（Raleigh North Caliana Star）的两位编辑 Thomas Henderson 和 Calvin Hones 鉴于北卡罗来纳州被明显地分为东西两区，两区的民众彼此接触很少，他们担心缺乏沟通会造成两区民众互不相识，从而扩大东西两区之间的差异。

在1810年3月30日，Thomas Henderson 和 Calvin Hones 进行了一项问卷调查，调查对象为北卡罗来纳州几个郡的居民，内容是有关农产品以及居民生活方面的情况。两位编辑把调查问卷发放和寄给他们所认识的人，以及被人推荐的居民。这次调查的结果没有刊登出来，不知道是因为什么原因。在新闻史文献中，这次调查被看做是新闻界所进行的第一次民意调查。

二、早期的选举研究

最早使用精确新闻报道技巧进行选举研究的是《哈里斯堡宾州人报》（Harrisburg Pennsylvanian），该报在1824年以模拟投票的方式，在维明顿市进行了美国总统选举的民意调查。结果显示，Andrew Jackson 共获得335票，John Quincy Adams 获得169票。调查结果刊登在1824年7月24日的《哈里斯堡宾州人报》报上。

在1851年《纳奇兹快讯报》（Natchez Courier）进行了密西西比州选举的民意调查，该报从密西西比州的行政区中，选出17个县，然后在每个县中，各访问数十位合格的选民，总计访问了800名选民。结果在17个县中，正确地预测出15个县的选举结果。

这些早期的精确新闻报道，在我们今天看来，方法上是很



粗糙的，并不符合科学研究的要求。但这些早期的精确新闻报道，显示新闻界的一种看法，即一般大众的意见和行为，可以成为重要的新闻。

到了20世纪初期，美国各大城市的报纸开始把大众的意见纳入例行的新闻报道中，层出不穷的各类选举活动，也促使更多的报纸来预测甚至解释选民的投票行为。

三、1936年的总统选举调查

1936年是美国的总统大选年，总统候选人分别是共和党候选人兰登（Alf Landon）和民主党候选人罗斯福（Franklin Roosevelt）。著名杂志《文学文摘》（Literary Digest）按照电话号码簿和汽车登记名簿上的地址，共寄出了1000万张模拟选票，回收了230多万张，结果显示，有57%的人会投兰登，支持罗斯福的比例为43%。于是，《文学文摘》根据调查结果，预测兰登将会在选举中击败罗斯福，当选美国总统。

在1936年，拥有电话、汽车的家庭都是比较少的，其中，拥有电话的家庭比例在33%左右[①]，绝大部分是中上阶层的民众。而选民不仅包括中上阶层的民众，也包括中下阶层的民众。所以《文学文摘》的调查对象对可能投票的选民的代表性很差。

盖洛普（Gallup）察觉到了这一点，于是，他采用代表性比较强的抽样方法（配额抽样），也进行了一次选民调查，结果显示罗斯福将会在大选中以56%的选票获胜。

大选结果揭晓，罗斯福获得了62%的选票，连任美国总

---

① Sheldon R. Gawiser & G. Evans Witt《A Journalist's Guide to Public Opinion Polls》Praeger Publishers，1994，第19页。

统。1936年的总统选举调查，《文学文摘》声誉大跌，而盖洛普名扬天下，使得科学的抽样方法受到重视。

从此以后，盖洛普以及其他的一些民意调查机构，如Crosslry、哈里斯（Harris）和罗珀（Roper）等，开始向新闻媒体出售民意调查的结果。大多数媒体认可民意调查的新闻价值，但没有自己去作调查，绝大多数的报社乐于刊登盖洛普、哈里斯等著名民意调查机构的调查结果。

四、1936年到1948年民意调的查发展

1936年到1948年期间，专业的民意调查活动发展迅速。在第二次世界大战前，美国农业部做了大量的调查项目，以确定如何提高抽样方法的效率。1937年，有关公众舆论研究的学术刊物《Public Opinion Quarterly》创刊，1941年全国民意研究中心（National Opinion Research Center，NORO）成立，现附属于芝加哥大学。

在第二次世界大战期间，受政府机构（如情报部门和陆军的研究机构）的委托，研究人员在美国国内和国外，进行了大量的调查活动。

在1940年和1944年的总统选举中，有关总统选举结果的预测依然正确。民众对调查活动越来越认可。

1946年，首次有关调查研究的学术和商业性会议在科罗拉多Central城（Central City，Colorado）召开，会议讨论了调查活动的发展，以及如何应对调查行业的发展。

1947年美国民意研究联合会（American Association for Public Opinion Research，AAPOR）成立，至今仍有影响力，非常活跃。



五、1948年的总统选举调查

1948年的总统选举调查对调查研究活动来说，既是一场灾难，又是-·个转折点。一场灾难是指多家调查公司对1948年的总统选举预测失败，使得民意调查的可信度受到怀疑；而另一方面，预测错误促使调查业更加重视科学的调查研究方法。

1948年总统选举的两位有竞争力的候选人分别是共和党托马斯杜威（Thomas E. Dewey）和民主党的杜鲁门（Harry S. Truman），后者是在罗斯福总统任内去世担任美国总统，谋求的是连任。

3家著名的民意调查机构盖洛普、罗珀、Grossley通过调查，都预测杜威将战胜杜鲁门，当选为美国总统。

而大选结果却正好相反，杜鲁门获得的选票比杜威多4%，连任美国总统。

具体的选举结果和调查预测结果如表1-3所示：[1]

表1-3

百分比	共和党 Dewey	民主党 Truman	独立候选人 Thurmond	独立候选人 Wallace	合　计
选举结果	45.1	49.5	2.4	2.4	99.4
Crossley	49.9	44.8	1.6	3.3	99.6
Gallup	49.5	44.5	2.0	4.0	100.0
Roper	52.2	37.1	5.2	4.3	98.8

① Norman M. Bradburn & Seymour Sudman《Polls & Surveys-Understanding what they tell us》，Jossey-Bass Publishers，1988年，第29页。

大众媒介过分相信了民意调查的结果，以至于大选结果出来后，令人大吃一惊，措手不及，在《芝加哥论坛报》（The Chicago Tribune）报道大选结果的版面上，登的是杜鲁门的获胜照片，而标题却是杜威获胜。

有关总统选举的调查大都是各大报纸委托调查机构实施的，预测失败，使得媒介大为恼火，一时对民意调查批评之声不绝于耳。很多人把这次调查业的失败与1936年《文学文摘》的失败联系起来，在1936年以前，按《文学文摘》的调查方法所进行的所有预测全部正确，而1948年之前，按现有调查方法所进行的所有预测都是正确的。

从事民意调查的人十分想搞清楚，到底是什么原因导致了预测失败？他们对研究方法进行了检讨。在1949年的AAPOR大会上，Crossley、盖洛普和罗玻设立了一个公开的论坛，讨论出现的问题。

社会科学研究协会（Social Science Research Council）对调查公司预测失败的原因进行了研究总结：

1. 调查结构停止调查的时间太早，忽略了那些最终转向杜威的选民；

2. 样本中高收入和中等收入的人比例过多，而这些人是倾向于共和党的；

3. 调查所使用的方法无法决定被访者是否会去投票；

4. 没有采用客观的、一致的方法把那些没有决定投谁一票的选民分配给各个候选人；

5. 简单随机的变化的偶然性使得样本的结果不同于总体的结果。

1948年以后，对选举的调查在方法上有了很大变化，表现如下：



1．民意调查一直持续在进行，直到选举日的当天，使得投票的趋势被看得很清楚；

2．抽样方法从松散的配额抽样，改变为修正的地域抽样，访问员进行入户访问；

3．对于判断是否投票，调查机构有了更多的方法。

1948年的选举预测失败，使得短期内人们对民意调查持怀疑态度，几家报纸停止刊登盖洛普的民意调查结果，同时对有关市场调查研究也提出了很多问题。但一段时间以后，民意调查和调查数据的使用者觉得他们还是需要通过调查得到的信息的，只不过在使用这些信息时要更加小心。

1948年以后，有关选举的调查预测结果绝大多数都是正确的，但还是存在着一定的误差，这使人们认识到，调查过程不可能是完美的。至少，随机的变化性就是原因之一。另外，还没有一个客观的方法来判断调查时没有决定投谁票的人，也没有客观的方法来判断一个被访者最终是否会去投票。最近几年的民意调查已经意识到这些问题，已经减低了关于不会出错误的可能性的声明。

六、1948年以后民意调查的发展情况

1948年以后，民意调查的发展十分迅速。应用民意调查的用户越来越多，应用范围也越来越广。

在1960年的总统选举中，总统候选人开始雇用了专业的民意调查人员/机构参与竞选活动，候选人可以利用调查来征求民意，从而制定竞选策略、主张、发展计划。约翰·肯尼迪（John F.Kennedy）雇用了哈利斯（Louis Harris），而理查德·尼克松（Richard Nixon）雇用了Claude Robinson。从此以后，每个总统候选人都雇用调查机构参与竞选活动，并且这种行为逐渐

扩展到州以及地方选举。一个调查机构，如果它的客户能竞选成功，特别是竞选总统成功，则可以为调查机构带来巨大的商业业务以及巨大的社会影响。吉米·卡特（Jimmy Carter）所雇用的调查人员哈利斯（Louis Harris）和肯德尔（Patrick Caddell）就是这样的例子。

由于社会上对调查的需要越来越多，调查机构也应运而生，越来越多。几个成立较早的调查机构发展壮大，一些州立大学，如伊利诺斯（Illinois）、威斯康辛（Wisconsin）以及加利福尼亚（California）都成立了专门的调查机构，从事本州以及地方上的调查。

1968年以及1972年两次美国总统大选期间，三大广播网ABC（美国广播公司）、CBS（哥伦比亚广播公司）、NBC（全国广播公司）开始委托调查机构进行民意调查。到1976年美国总统选举时，精确新闻报道已经蔚然成风，不论是全国的媒介，还是地方性的媒介，都开始自己进行民意调查，精确新闻报道成为选举新闻的最重要的特色之一。此后，精确新闻报道发展迅速，三大广播网和规模大的报纸定期进行精确新闻报道研究，一些地方性小报偶尔为之。

在1978年，有学者（Rippey）对437家日报进行了访问，发现其中有162家已经进行过民意调查[①]。一般认为，20世纪70年代，精确新闻报道已经成为美国新闻界普遍重视的一种新闻报道方式。

七、菲利普·迈耶的贡献

菲利普·迈耶（Philip Meyer）1948年至1952年就读于坎萨

① 罗国辉《精确新闻报道》，正中书局，第35页。



斯州立大学（Kansas State University），1956 年至 1958 年在北卡罗来纳大学（University of North Carolina）学习，获政治学硕士。曾在多家报社当过记者，后在大学任教[①]。

在对当时流行的新闻学的手法进行了研究以后，迈耶认为利用诸如描写入微的细节、新闻人物的内心独白以及由人物、冲突、结局组成的短篇小说结构等文学表现手法，新闻记者可以写出引人入胜的作品。但是，要使新闻读起来像小说，记者必须收集大量的素材，从中挑选自己合适的，而这一点对于新闻报道来说是很难做到的，因为新闻要求的是真实性，时效性要强。迈耶认为将文学手法引入新闻报道，本质上是将新闻推向了艺术，而艺术并不适合有着多种严格限制的新闻报道。理想的解决办法是让新闻与科学相结合，迈耶主张将社会科学与行为科学的研究方法引入新闻报道。

1967 年迈耶担任骑士报业集团（Knight-Ridder）的国内通讯员，底特律市爆发了严重的黑人暴动骚乱，迈耶和另外两位社会学家，采用随机抽样的方法，对居住在底特律种族骚乱地区的 437 位黑人进行了一次访问调查，把结果输入了电脑，分析结果显示，参加骚乱的黑人中受教育程度没有显著的差异。当时流行的观点认为，骚乱的根源在于在南方长大的黑人难以与北方的文化融合，分析结果表明，参与骚乱的人大部分是在北方长大的。

在对调查数据进行分析的基础上，迈耶写出了《十二街那边的人们》，发表于《底特律自由报》（Detroit Free Press），并把这种报道方式称为精确新闻学。

迈耶对底特律黑人骚乱的系列报道，不仅为《底特律自由

---

① 摘译自 http: //www.unc.edu/ ~ pmeyer/meyeres.htm.

报》赢得了1968年的普利策新闻奖（Pulitzer Prize），也使得精确新闻报道名声鹊起，逐渐受到美国以及世界各国新闻界的重视。

1973年菲利普·迈耶出版了《精确新闻学——一个记者关于社会科学方法的介绍》（Precision Journalism: A Reporter's Introduction to Social Science Methods）一书。在很多的新闻学课堂上和新闻报道的高级研讨班上，该书成为标准的参考手册。迈耶原想把书名叫做《社会科学方法在新闻学实践中的应用》（The Application of Social Science Methods to the Practice of Journalism），但一位新闻学教授Everette Dennis建议叫做精确新闻学，迈耶采纳了这个建议[①]。

在1979年《精确新闻学——一个记者关于社会科学方法的介绍》出版了第二版。1996年美国民意研究协会（AAPOR, American Association for Public Opinion Research）把该书列为关于舆论研究最重要的50本书之一；2000年，《新闻学季刊》(Journalism Quarterly) 把该书列为20世纪新闻学与大众传播学领域最重要的35本书之一[②]。

有关菲利普·迈耶的详细资料，请参见http://www.unc.edu/~pmeyer/meyeres.htm。

八、新闻教育的影响

美国新闻教育的改革也间接促进了精确新闻报道的兴起。

在1930年以前，美国大学的新闻系或新闻学院隶属于文学院，新闻教育侧重新闻业务训练，开设的课程以新闻报道、

① David Pearce Demers与Suzanne Nichols《precision journalism-A practical Guide》，第12页。

② 摘译自http://www.unc.edu/~pmeyer/meyeres.htm。



消息写作、社论写作以及新闻史为主。

在20世纪30年代初期，威斯康辛大学率先在政治系以及社会学系的博士课程中，增加了新闻课程。博士课程重视研究方法与理论的训练，这两个系毕业的经过严格的社会科学方法与理论训练出来的博士，把所学的社会科学的方法与理论带到了美国最著名的新闻学院，如史坦福大学、明尼苏达大学、西北大学等，使得这几所在学术界居领先地位的新闻学院，开始重视社会科学教育。

到了20世纪50年代，美国主要大学的新闻学院纷纷成立博士班，除了传统的新闻业务的训练以外，重视社会科学研究方法与理论的教学。到了20世纪70年代以后，美国各类大学的新闻学院大都开设了研究方法方面的课程，有的学校甚至把研究方法纳入采访写作的课程中。在1978年有学者对全美77所设有新闻硕士课程的大学进行了调查，发现61%的学校把研究方法列为必修课[①]。

由于新闻学院越来越重视研究方法，所以毕业生中越来越多的人能够执行精确新闻报道，这些人到了各个媒体以后，又促使新闻机构重视报道精确新闻。

九、电话调查

近40年来，民意调查的巨大变化是得益于科学技术的发展，特别是电话的普及和电子计算机的广泛使用。

在民意调查收集数据的过程中，主要有3种方式，面访调查、邮寄调查以及电话调查，电话调查比起面访调查和邮寄调查，成本降低，周期缩短，更能适应精确新闻的需要。随着经

① 罗国辉著《精确新闻报道》，正中书局，第37页。

济的发展，居民的家庭电话普及率越来越高，使得电话调查成为可能。

目前，在美国，家庭电话普及率已经超过了95%，这样的话，就可以通过电话调查的结果推断全国民众的情况。

十、电脑及统计软件的影响

由于精确新闻报道的资料分析工作十分繁重，单靠手工操作往往费时过长，甚至是难以胜任的。

早在1952年，新闻界有过一次在新闻报道中使用计算机的尝试。在1952年11月4日总统大选揭晓的当天，哥伦比亚广播公司决定在电视实况转播中，利用计算机，根据早期得到的部分选举结果对最终结果进行预测。预测结果是艾森豪威尔将以压倒票数当选美国总统，但在场的主持人以及有关人士对此结果不敢相信，所以没敢播报。直到午夜，大选结果出来，才证明利用计算机预测的结果是正确的。从此以后，通过计算机利用有关的统计软件、根据部分选票的结果对最终选举结果进行预测，已经成为大选揭晓之夜电视实况转播的一个组成部分。

在20世纪70年代末、80年代初，电脑在新闻媒介还没有得到广泛的应用。极少数大规模的新闻机构即使拥有一两台电脑，也都是用来进行人事档案或财会管理的。当时使用的计算机大都结构复杂，成本昂贵，要想让计算机完成某项运算，必须由专职的程序员来编写命令，人工代价很高。另外，当时很少有新闻记者懂得用计算机进行数据分析，相当一部分的记者对计算机心怀畏惧。这一切使得新闻媒介对利用计算机进行资料的统计分析持谨慎态度。

20世纪80年代，微型计算机开始在市场上出现，具有里



程碑意义的是 IBM 微机，IBM 公司将其电脑的设计和结构公开，任何制造商都可以加以仿制，而不需要许可证协议，结果导致众多厂家生产出大批可与 IBM 兼容的微机，批量生产降低了成本，使得电脑价格迅速下跌，微电脑行业发展势头迅猛。

与此同时，软件行业也进入了蓬勃发展阶段，不断开发出界面友好、方便使用的各类软件，包括文字处理、数据统计分析等。目前，用来进行社会科学资料分析的软件有很多种，常用的有社会科学统计软件包 SPSS（Statistics Package for Social Science）、SAS（Statistical Analysis System）、STATA、SYSTAT、SURVTAB 等等。

对于新闻界来说，电脑化浪潮不仅使得办公自动化，而且也使得在短时间内分析大量数据资料成为可能，推进了精确新闻报道的发展。

十一、电脑辅助的电话访问系统 CATI

电脑和电话结合起来，就产生了电脑辅助的电话调查（Computer-Assisted Telephone Interviewing）。简单说来，就是按照要求，电脑自动进行抽样，拨打电话号码，访问员工作在电脑屏幕前，按屏幕的提示来提问，直接录入答案，并且电脑可以即时进行数据的处理分析。

电脑辅助的电话调查使得调查访问的周期更短，使得精确新闻的时效性更强。是精确新闻采集数据的主要方式。

有关电脑辅助的电话调查，是本书的重点，我们在第七章中专门来介绍。

第三节 美国主要媒体从事民意调查的情况

美国三大广播网 ABC、NBC 和 CBS 都建立有自己的调查访问中心。

美国 CBS 从 1967 年开始，NBC 从 1973 年开始从事民意调查活动，把调查结果在晚间新闻以及其他的专题节目中播出。1981 年 ABC 也建立自己的民意调查中心。三大广播网都有一个印刷媒体作为合作伙伴，帮助设计调查并分担费用。CBS 的合作伙伴是《纽约时报》（New York Times），ABC 同《华盛顿邮报》（Washington Post）合作，NBC 在 1978 年至 1983 年是同美联社（Associated Press）合作，以后是和《华尔街日报》（Wall Street Journal）合作。

一、调查的主题

调查涉及的主题包括总统支持率，消费态度以及有关政治、经济的其他话题。三大广播网民意调查的主题详细如表 1-4。[1]

二、调查的覆盖范围

从调查覆盖范围来看，包括了除阿拉斯加和夏威夷以外的所有美国大陆地区。

---

① Norman M. Bradburn & Seymour Sudman《Polls & Surveys-Understanding what they tell us》，Jossey-Bass Publishers，1988 年，第 127 页。



表 1-4

	主　　题
全部包括	流产、公民防卫、犯罪、古巴、死刑、经济、教育、选举、萨尔瓦多、福克兰（马尔维纳斯）群岛、枪支控制、移民、国际关系、劳力、中东、党派识别、波兰、总统支持度、性/色情文学、社会保障、苏联、税
CBS & ABC	空中交通管制、对美国制度的信任、美国所面临的重大问题、种族态度
CBS & NBC	平衡预算、篮球罢工、国会、征兵、伊朗、石油、空间计划、荒地政策、妇女
ABC & NBC	道德优势群体
CBS	阿富汗、空调装置、原子能、中国、能源、环境、医疗保险、墨西哥、奥运会、巴拿马运河、旅行
ABC	AWACS、拳击、过度防御、侵犯隐私权、南希·里根、电视新闻采访
NBC	居住、最高法官提名、再提水门事件、荒地政策

三、调查访问的方式

从收集数据或者说访问方式上来看，三大广播网的民意调查都采用电话调查的方式进行。电话调查被采用，原因有二，一是面访调查成本越来越高，二是美国的家庭电话普及率已达到 95%以上。

四、抽样方法

抽样是分为 3 个步骤来进行的。

1. 抽取电话局号

从所有区号和电话局号的抽取一部分。局号的全部名录是通过有关公司得到的。

2. 抽取电话号码

局号确定后，再加上4位数字就可以构成一个电话号码，但是按这种方法构成的号码只有1/5的可能是住宅电话，有4/5的可能是没有投入使用或者商业、学校以及其他的非住宅电话号码。为了减少拨号的费用，调查机构尝试着判断出无效的电话号码。ABC有一个包含所有公开号码的文件，从中可以排除没有投入使用的号码以及非住宅电话号码，然后进行抽样。

CBS采用尾数加1法产生电话号码。

3. 抽取访问对象

研究表明，在住宅电话响铃以后，在成年人中，女性接听电话的机会是男性的3倍，长者接听电话的机会多过年轻人。①

抽取被访对象多抽取最近一个过生日的人。但这样获得的样本常常男性偏少，因为男性倾向于拒绝接受访问。可以采用先要求访问一个男性，如果没有男性的话，再访问女性。

事实上，这样所获得的样本不是自加权的，需要事后加权。CBS利用地区、种族、性别、年龄和文化程度进行加权。所选择的加权变量要与所询问的问题密切相关。

五、影响访问的因素

调查的内容和长度会影响被访者的合作，被访者对调查内

① Norman M. Bradburn & Seymour Sudman《Polls & Surveys-Understanding what they tell us》, Jossey-Bass Publishers, 1988年，第127页。



容越感兴趣，合作的比例就高。研究发现，有关健康、家庭的话题比外国政治更受欢迎。一般说来，电话访问都比较短。

六、样本量

从访问的人数上来看，每个调查项目的样本量一般在1200～1500之间。

第四节　精确新闻在中国的发展

精确新闻报道在中国的发展是近20年的事情。1983年1月29日英文版《中国日报》（China Daily）在国内新闻版以头条位置报道了北京新闻学会在1982年完成的北京市受众（读者、听众、观众）调查的结果，标题为《首次运用电子计算机进行民意测验表明：读者相信报纸》（First computerized opinion poll shows support for press）。

北京市读者、听众、观众调查是由北京新闻学会发起，中国社会科学院新闻研究所、《人民日报》、《工人日报》、《中国青年报》、北京广播学院参与的一次抽样调查。

调查目的是了解读者、听众和观众的需求，为报纸、广播和电视的改革提供依据。调查时间为1982年6月8日至8月6日，调查总体为北京市13岁以上的市民（不包括城市无职业的家务劳动者），调查采用统一问卷，直接访问的方式进行，共回收有效问卷2423份[①]。

此后，类似的关于调查结果的报道陆续在一些报纸中出

① 北京新闻学会调查组《北京读者、听众、观众调查》，工人出版社，1985年1月，第1～3页。

现，但只是以片断的形式零散地出现，并且大多属于调查机构针对社会热点问题进行调查，为了宣传的目的，把调查结果提供给媒体，以提高调查机构的知名度。由媒体自己来组织或委托的民意调查非常少。

进入90年代，新闻改革的深入，推动了媒体进行精确新闻报道。经济的发展，使得社会生活的节奏加快，人们更需要看到对社会生活全景结构式的报道。而以民意调查为代表的精确新闻报道作为一种科学可靠的、对于社会事物和社会现象进行系统化观察的手段，由于其独特的功能、特长而迅速地被人们发掘并启用了。

1993年1月《中国青年报》社会调查中心成立，任务就是专门就人们所关心的社会热点问题进行有关的民意测验，并且每月两期开辟版面专门用来报道有关的民意调查的结果。

1994年1月1日《北京青年报》在全国报刊界率先以每周一期的频率推出《公众调查版》，声称“用科学的眼光观察世界，给你一个量化的真实”。

同年，中央电视台每周一期的《与你同行》节目开设《微型调查》栏目，用数据化的统计方法结合电视的形象性特点进行精确新闻的报道[①]。

1996年1月3日《北京青年报》新闻周刊在其公众调查专版创办两周年之际，将其公众调查专版冠以“精确新闻”的头衔，精确新闻报道被我国媒体所认知。拉开了我国媒体科学、规范的精确新闻报道的序幕。

1997年8月，中央电视台经济部同国家统计局中国经济景气监测中心合作，在《中国财经报道》的周末节目中推出独

① 喻国明《解构民意》，华夏出版社，2001年，第22页。



家新闻报道《每周新闻》，结合经济部的特点，把报道综合经济内容作为节目的出发点。在开播的头一年中，播出了近70期节目，既有涉及国家大局的宏观经济题材，也有反映行业、部门经济问题的题材，还有分析个人消费和投资行为的微观经济题材，有专家认为，报道内容都是市场经济下，群众所关注的热点、难点、疑点问题[①]。

据不完全统计，目前，国内已有几十家报刊定期或不定期地刊登精确新闻，越来越多的媒介已经直接开办了相关的专栏。表1-5列出了媒介名称和所开设的报道精确新闻的栏目名称。[②]

表1-5

媒介名称	栏目名称
《南方周末》	数据
《投资导报》	数字视点、财源数据
《粤港信息日报》	今周图表
《中国消费者报》	数据·生活、货比三家、比较价格
《中华工商时报》	商情调查、地方调查
《北京青年报》	精确新闻
《中国经营报》	微型数据、市场调查
《中国青年报》	社会报告、数据库、微型调查
《文汇报》	调查

① 国家统计局中国经济景气监测中心，中央电视台经济部《中国财经报道》《中国都市生活报告》，中国统计出版社，1999年5月，第303页。

② 李文刚：《用数字说话—数字新闻的理论与操作》，天津人民出版社，2000年11月，第66-67页。

续表

媒介名称	栏目名称
《深圳商报》	与我们生活相关的数据
《经济日报》	景气观察
《景气运行月报》	分类景气图
《中国信息报》	观海潮
《西北信息报》	数据库
《中国汽车报》	社会调查
《北方周末》	数字生存
《中央电视台》	舆论调查、CCTV 网络调查、新浪网调查

第五节 报道精确新闻的程序

在科学研究中，不论是自然科学还是社会科学，都要遵循一定的程序进行。精确新闻报道采用的社会科学的研究方法来收集材料，报道新闻，也要按照一定的程序来工作，目的是避免不必要的错误，提高新闻报道的质量。

一般说来，精确新闻的报道程序是这样的：

一、选择报道题材

选择精确新闻的报道题材时，可以从现有的报纸杂志、学术期刊、学术研讨会以及有关的历史档案中获取灵感。

并不是每个题材都适合做精确新闻报道的。进行精确新闻报道一般说来费时费力，费用也比较贵，所以在找到一个自己感兴趣的题材后，要考虑是否具有新闻价值、是否适合进行实



证研究、题材大小是否合适等因素，以决定这个题材是否适合进行精确新闻报道。

二、熟悉报道题材

研究题材确定以后，记者可以通过查找资料，专家咨询等途径，来了解是否有人进行过类似的研究？有哪些资料可供参考？所要进行的精确新闻的研究，涉及到变量是什么？这些变量的操作化定义是什么？

三、提出问题或假设

下一步就是提出研究问题与假设，并确定好研究方案，包括采用什么方式进行资料收集，是抽样调查还是内容分析？如果采用调查的方法来收集资料的话，研究设计要包括问卷设计、确定抽样方法、确定访问方法等等，如果采用内容分析的话，要确定抽样方法、类目建构等等。

四、资料收集

研究方案确定后，就可以开始数据收集工作，包括复核、检查等。完成以后，一般要进行数据的录入、整理。

五、资料分析与解释

数据整理好以后，就可以开始统计分析工作。

六、新闻写作

根据研究问题以及研究假设，结合数据分析结果，进行新闻写作。



应 用 实 例

以下为获1968年普里策新闻奖的关于底特律骚乱的系列报道的一部分。[①]

底特律骚乱，追查43人死亡原因

——报社全体出动，报道该市最大惨闻。

1967年夏天，美国的种族冲突达到了历史上从未有过的最高潮。

继纽瓦克市和纽约市哈莱姆区骚乱后，底特律这个火药桶也爆炸了。

从7月23日到30日之间，在大部分黑人区，烧、杀、抢频繁不断。有43人死亡，成百上千的人受伤，五到七千人无家可归。

政府出动了4700名陆军伞兵部队士兵和8000名国民警卫队员在该市大街上值勤，维持治安。

底特律的两家主要报纸《自由新闻报》和《新闻报》史无前例地投入全部人马，动用一切消息来源、设备和资金来报道这一事件。报社的记者、撰稿人、摄影记者、编辑每天都只休息几个小时，通宵达旦、废寝忘食地投入报道活动。

就在这场骚乱尚未平息前，《自由新闻报》的编辑决定立即对此次事件进行一次大规模的调查，访问受害者、警方、部队和报社的工作人员，探讨事件发生的原因。

① 摘自［美］约翰·霍恩伯格编著《普利策新闻奖获奖作品选》，第131～135页。



在骚乱彻底平息后，报社又派出记者前往调查。他们奋战了5周，每周工作7天，每天不少于12小时，写成了一篇2.4万字的报道，在报上占了5个版面。以下是报上发表的结论部分的摘要：

为什么底特律会有43人死亡?

（选自1967年9月3日《底特律自由新闻报》）

在8个沉寂的白昼和8个枪声不绝的夜晚，底特律大街上，有43人死亡或身负重伤。他们被解释为骚乱的牺牲品，他们是现代内战的受害者。

即使现在，也很难作出比较令人满意的解释。最吸引人的结论是，只有骚乱本身才应对这种局面负责，在这种情况下，死人是极其平常的事。

就在骚乱开始36小时后，政府出动了3000多名全副武装的人员，在靠近西区的每一个街区有20名武装人员值勤，集中这么多的兵力只有大规模军事入侵才能与之相比。

在底特律东区，驻有成百上千正规陆军伞兵部队。在市区、市里和州里的警察部队以及国民警卫队员乘坐巡逻车、警车、快速巡逻车、吉普车、卡车、装甲运兵车和坦克到处巡逻。武器包括值勤左轮手枪、M-1型步枪、私人运动枪、短筒连发手枪、卡宾枪和50毫米口径机关枪等。

这许多武器必然会引起对抗，会出事，会发生差错，最后，会有人送命。武器和军人太多，必然导致这样的结果。

现在，虽然还没有答案，关键问题却很简单。

——43人中有多少人应死去?

——其中有多少人可以避免死亡?

《底特律自由新闻报》派出了一个记者组对每一名死者都

进行了独立调查，在此基础上，对这些问题一一作出了解答。

调查得出这样的结论：

大多数牺牲者不应死去。他们的死亡是可以避免的，而且应该避免的。

命运对死者的安排是没有什么条条框框可遵循的，而骚乱的牺牲者也不是可任意归类或分成三六九等的。其中，最为无辜的是一个4岁的小女孩，她是被流弹打死的；最该死、罪有应得的是一个喝得烂醉的冷枪手，他是在朝别人开枪时，被打死的。每个人死去的方式都不太一样。

有18名牺牲者是底特律的警察开枪打死的，其中有14名，在《底特律自由新闻报》调查后，被认为是抢劫犯。另外4名是枪手，有1名可能是纵火犯，但未经证实，其中有两人是在阿尔吉尔汽车旅馆被打死的。

至少有6名牺牲者是国民警卫队打死的，其中有5人完全无辜，现在看来，他们是死于悲剧性的事件。

还有5起人命涉及到警察和国民警卫队，很难确定是谁的子弹使这些人丧生。

还有两人，都是抢劫犯，是被商店的主人开枪打死的。还有3人是被普通居民打死的；两起已下达了逮捕证，另一起，正在考虑下达逮捕证。

有两名抢劫犯在一家商店偷东西时，被卷到大火中烧死了。

还有两起死亡事件，一是消防队员，一是普通市民，他们触电而死。

最后剩下5起死亡事件：

一名19岁的男青年被一名伞兵误伤，中弹而死；

一名23岁的白人妇女被身份不明的枪手打死；



一名底特律市消防队员可能被暗藏的枪手射死，也可能被国民警卫队的流弹射中；

一名警察与囚徒搏斗时，旁边一名警察被打死；

阿尔吉尔汽车旅馆还有一人死亡，凶手不知去向。

事后议论是非曲直，是很容易的。火焰已经扑灭，大街上十分平静。人们认为，在骚乱过程中，本来应当有理智、秩序和通情达理，但那是不可能的。

当火焰还在继续蔓延，有些影子还在空荡荡的走廊上徘徊时，当"站住"的叫喊声划破寂静的夜空，吓人一跳时，当脚步在破碎的玻璃上踩出刺耳的声音时，任何在半夜乘车经过大街的人都会认为，在这个恐怖的时刻，人们的行为不会是理智的。

以下是对骚乱死者进行5周调查后，根据上述可以接受和理解的情况，得出的总的结论：

——对在骚乱地区活动的枪手的数目及其危险性作了过高的估计。死者中只有1人是枪手，只有3名受害者可能是枪手打死的，有两起还有疑问。有30人因指控为枪手被捕，但还没有一起案件受到审理。

——在43起死亡事件中，只有7起可能是故意杀人：阿尔吉尔汽车旅馆的3起，居民开枪打死的3起，还有1起，即威廉·达尔顿案件，还没有头绪。对其余的死亡事件，《底特律自由新闻报》还没有找到任何有意或预谋杀害的证据。

——现在回想起来，密执安州和底特律市的警方一般说来还是克制的，给人留下深刻的印象。事实上，4400名警察在动乱之中度过了至少5天时间，流血事件并不算多，这也就不错了。当然，也有个别情况判断失误。这些令人遗憾的事件可能是他们开枪太快，虽然他们的行动是合法的。

这里提出一点主要批评。底特律市当局和军方都想将死亡的人数控制到最小的程度，虽然他们采取的方式各有不同。但是，他们的努力并不成功，造成了一些不必要的死亡。

星期一晚上11点20分，就在国民警卫队归联邦当局控制后数小时，指挥官约翰·思罗克莫顿中将对其统帅的所有部队下了一道总命令，命令士兵卸掉武器中的子弹，只有在军官的指挥下才能开枪。

思罗克莫顿中将的正规部队遵从了这一命令。命令下达后5天内，在伞兵的值勤区只有1人被打死。

国民警卫队在多数情况下都没有执行这一命令，因为命令没有通过正常的途径传达下来，在大街上巡逻的国民警卫队员对此根本就不清楚。其结果是，国民警卫队开枪打死了11人，其中有9人完全无辜。

如果军纪严明，执行思罗克莫顿中将的命令，这些死亡是完全可以避免的。



第二章　精确新闻学的基本概念

精确新闻学的主要特色是采用社会科学的方法进行新闻采访，本章所介绍的有关精确新闻学的一些基本概念就是社会科学研究方法的基本概念。

在社会科学研究中，前提假设是：整个世界是有秩序的；存在个人规律与社会规律；研究中的行为规律是可以观察和测量的。在这些前提下，社会科学研究方法的主要特征是客观和系统。客观是指研究结果不能掺杂研究者的个人喜好，也不能让客户对结果的喜好主导研究工作。系统是指研究不是随意选择的，不能为取得某一结果而有意操纵研究。

与自然科学相比，社会科学的研究对象不是无生命的物体或不会深思熟虑的动物，而是具有自我意识的人。人有着复杂的个性，并能由于合理的或不合理的原因选择自己的行动方针。社会学家研究的是人，这一事实为社会科学研究方法论提出了5个困难：

第一，调查社会行为的做法本身就有可能改变被调查者的行为。当人们知道有人在研究他们的时候，他们的行为也许会一反常态。

第二，人有感情、有动机、有其他高度的个性特征。人可

能由于错误地领会了一个问题或者由于不懂得他们本身的行为或态度产生的原因而有意无意地提供错误的信息来突出自己，还可能由于本身的种种特殊原因而干出出人意料的事情来，而这些，无生命的物体是不会做的。带来的结果是，社会科学的解释和预测常常不如自然科学那样准确。

第三，造成社会行为的原因几乎历来是极其复杂的，涉及到许多社会的、心理的、历史的和其他的因素。社会科学通常比自然科学更难分清因果，因为这里涉及到的变量往往要多得多。要确定水为什么会沸腾或者是压力对某种气体体积的影响是比较容易的。但要确定人为什么会恋爱或者为什么会爱上所爱之人，就难得多了。

第四，由于伦理学上的原因，有些实验不允许在人身上进行。自然科学家用光线做实验时没有任何道德上的顾忌，生物学家用动物做实验时也很少感觉内疚。但是，人的尊严和隐私权却必须受到尊重。我们不能故意把男孩当成女孩养大以便看一看这种经验对孩子今后适应性别角色有什么影响，无论这项研究的成果可能是多么有趣和有价值，我们也不能这样做，伦理学方面的考虑严格限制了社会科学家可以使用的方法。

第五，与自然科学家不同，社会科学家本人也正是自己的研究对象的一部分，因此，要想保持一种超然的态度，就要困难得多。地质学家也许对于确定某种岩石样本的成分很感兴趣，但他与研究成果不可能有什么感情上的纠缠。社会科学家也许会研究种族关系或贫困问题，他们在感情上与研究的结果纠缠在一起。研究者对于研究对象的困难或经历可能会产生强烈的共鸣，结果，调查过程和做出的解释就有可能失之偏颇。

所有的社会学家都承认这些问题，对于解决这些问题，多数人接受韦伯的观点，即社会科学的研究方法应该尽可能地模



仿自然科学，但是，由于研究的问题与自然科学大不相同，所以有时也需要解释性的、主观的方法。

第一节　研究的类型

按照不同的标准来分，社会科学研究可以分为不同的类型①。从研究目的来看，可分为探索性研究、描述性的研究以及解释性研究；根据研究涉及时间的不同，可分为横向研究和纵向研究；按照数据收集方式的不同，可分为定量研究和定性研究。在精确新闻学中，所涉及到的研究都是定量研究，以描述为主要目的，既有横向研究，也有纵向研究。

表 2-1

研究目的	时间	数据收集技术
探索性研究 描述性研究 解释性研究	横向研究 纵向研究： ■固定样组 ■趋势研究 ■世代研究	定量研究： ■实验 ■调查 ■内容分析 定性研究： ■田野研究 ■历史比较研究

一、按照研究目的分类

社会科学研究中，研究者对各种特定现象进行研究时，会

① W. Lawrence Neuman《Social Research Methods-Qualitative and Quantitative Approaches》，第三版，第 33 页。

有着各不相同的目的。但是所有社会科学研究的目的大致分为三大类，即探索性研究、描述性研究以及解释性研究。

不同目的的研究所对应的设计要求、研究对象和研究方法的选择，以及具体的操作程序都会有所不同。

1. 探索性研究

探索性研究（exploration research）一般是指研究者对所感兴趣的问题或现象进行初步的考察或了解。当研究者本人对打算研究的问题不太熟悉的时候，当所研究问题本身鲜为人知，很少有人涉及的时候，就经常采用探索性研究。

探索性研究可以使研究人员对问题或现象有初步了解，还可以为更深入系统的研究提供指导和线索，来确定是否进行正式的研究。

探索性研究对精确新闻报道意义重大。当记者对研究题材不太了解，或者面对着从未发生过的社会现象时，在正式的研究开始前，先进行探索性研究，可以帮助更好地掌握研究的重点。

一般说来，探索性研究的样本缺少代表性，使研究结果的代表性受到限制。因此，探索性研究只能对研究问题提出暂时性的答案。

2. 描述性研究

描述性研究（description research）是指研究描述社会事件与现象，如社会环境、资源、组织、人口特征以及民众对某一事件的认识、意见、态度与行为。描述性研究在对社会现象的认识上，比探索性研究前进了一大步。

描述性研究的目的在于描述事件和现象，可以解答社会现象“是什么”的问题，能对现象的状况、特点和发展过程做出客观、准确的描述。但描述性研究并不解释产生这些问题和现



象的原因。

系统周密的描述是正确认识和解释社会现象的前提，只有在客观真实地描述现状的基础上才能说明现象间的因果联系。此外，通过描述性研究还能发现一些新的现象和问题。

如我国每 10 年进行一次的人口普查就是典型的描述性研究。人口普查的目标就是对中国人口的各种基本特征进行准确的描述，从而提供中国人口在年龄、性别、职业、文化程度、民族、婚姻状况等方面基本情况。人口普查并不解释为什么中国的人口会有这些特征。

用于了解消费行为的市场调查也多是描述性研究，通常只是描述消费者对某个产品的使用情况，并不深入探讨某产品为什么受欢迎或不受欢迎。社会学领域的有关犯罪问题和婚姻问题的调查，也是描述性研究。

就精确新闻报道而言，常常用到民意测验的结果，基本上也是描述性研究，只是描述民众对某个事件或人物的看法，一般并不深入解释为什么会这样。

3. 解释性研究

解释性研究（explanation research）不仅描述社会事件与现象，并对产生的原因加以解释。

由于解释性研究的目标是回答“为什么”，因此理论色彩浓厚。通常是从理论假设出发，经过实地调查、收集资料并通过对资料的分析来检验假设，最终达到对社会现象进行理论解释的目的。

社会科学的许多领域的研究都采用解释性调查研究的方式。在调查方案的设计和调查程序上，解释性研究比描述性研究更为复杂、严谨，需要像自然科学研究那样，事先制定周密的实施方案。

和解释相联系并附于解释之上的目标是预测。预测是要揭示将会发生的事件和情况。有人认为，预测是比解释更高的要求，也有人认为解释和预测从根本上说是相同的现象，只不过预测在事件发生之前，而解释在事件发生之后。对于社会现象的研究，不仅可以在其发生之后做出解释，而且可以在解释的基础上对有关现象今后的发展做出预测，甚至在一定的条件下控制它们的产生。

解释性研究涉及理论假设，并需要对假设进行验证，或提出理论来解释研究发现，过程比较复杂，因此时效性比较差，可读性不强。精确新闻报道很少采用解释性研究。

总的说来，探索性研究、描述性研究和解释性研究概括了研究的不同目的，有些研究同时包含多种目的。

风笑天在《社会学研究方法》中给出了3种不同目的的研究类型特征比较，如表2-2：[①]

表2-2

	探索性研究	描述性研究	解释性研究
对象规模	小样本	大样本	中样本
抽样方法	非随机抽样	简单随机，按比例分层	不按比例分层
研究方式	观察、无结构访问	问卷调查、结构式访问	调查、实验等
分析方法	主观的、定性的	定量的、描述统计	相关与因果分析
主要目的	形成概念和初步印象	描述总体状况和分布特征	定量关系和理论检验
基本特征	设计简单、形式自由	内容广泛、规模很大	设计复杂、理论性强

① 风笑天《社会学研究方法》，中国人民大学出版社，2001年3月，第70页。



二、按照时间维度分类

按照社会科学研究所涉及的时间维度的不同，研究项目可以分为横向研究和纵向研究。

1. 横向研究

横向研究（Cross-sectional studies），又称横剖研究，一次性研究。只涉及一个时间段，横向研究的目的是对某种社会现象或某一社会总体的横截面进行了解。

如果研究对象是人的话，那么横向研究一般要广泛抽取该时间内各种各样的个人做样本，如不同年龄的、不同教育程度的、不同收入水平和不同民族的、宗教的等等，从而可以对不同类型的人的情况进行比较。

要注意的是，这里所说的时间段，根据具体研究项目的不同，可能是几天、几个星期或几个月。不管时间周期的长短，数据收集的工作只进行一次。

在民意调查中，属于横向研究的项目很多，横向研究的主要优点是能在同一时间从大量的人群中收集起资料，这些资料由于不受时间变化的影响而有很大可比性。

2. 纵向研究

纵向研究（Longitudinal studies），又称纵贯研究，所收集的资料要涉及数周、数月或好多年。纵向研究对研究对象持续进行研究，而不是一次性的。

纵向研究又可以细分为固定样组研究、趋势研究以及世代研究。

(1) 固定样组研究

固定样组研究（panel studies）也叫重访法，是指在不同的时间段对同一组研究对象进行了多次访问。

1997年7月1日香港回归祖国，在此事件前后，香港的一些学者进行了学术研究，采用了固定样组的调查方式。在香港回归前，即1996年的9月对广州1200位市民进行了问卷调查，内容是关于对香港的了解以及对香港的评价。在香港回归以后的1998年3月，采用同样的问卷对一年前调查过1200人再次进行了访问。这就是一个典型的固定样组调查。必须注意的是，随着时间的变化，第一次被访问的一些人，已经搬离了原来的住处，很难做到对原来的1200人都进行访问。

在广播电视的收视率、收听率调查中，抽选的样本户要持续一年左右才被换掉，也可以看做是固定样组。

（2）趋势研究

趋势研究（trend studies），也有人称时间序列研究（time series studies），是指在不同的时间段里，对不同的人，访问同样的内容。

目前十分流行的消费者满意度调查、消费者信心指数等采用的就是趋势研究。比如说，由国家统计局中国经济景气检测中心负责的消费者信心指数调查，是每个月进行一次，每次在北京、上海、广州、成都、武汉、西安和沈阳7个城市中随机抽取600位15岁至64岁的城市居民，以调查问卷的方式进行访问。调查问卷的内容主要包括5个方面：对当前经济形势的判断、对家庭收入的看法、对目前购买商品时机的判断、对未来整体经济的判断以及对自身收入的评判。

再如中国人民银行进行的城镇储户调查，是每个季度进行一次，在全国共抽取了50个城市，每个城市抽取了8个储蓄所，每个储蓄所每次调查50名储户。调查采用自填问卷的方式，内容涉及储蓄、消费等情况。每季度公布调查结果。

趋势研究关注的研究问题的发展变化的趋势和规律。



(3) 世代研究

世代研究（cohort studies），也叫同期群研究或人口特征组研究，是指研究某一特殊人群随时间推移而发生变化的情况。在世代研究中，每次的研究对象是不相同的，但所有研究对象有一个特征是共同的。

举例来说，在我国，在 1968 年上山下乡的“老三届”学生是人们关注的对象，如果以“老三届”学生为研究对象，了解他们在 1978 年、1988 年和 1998 年的情况，每一次可以调查不同的人，但这些人必须是“老三届”。

在美国，有社会学家针对那些在 20 世纪 30 年代出生的人做过世代研究，研究是每 10 年进行一次，1950 年，研究对象的年龄在 20 岁至 25 岁，1960 年，抽取的研究对象是 30 岁到 35 岁的人，1970 年，研究对象已经是 40 岁到 45 岁了。虽然 3 次研究多抽取的研究对象是不同的人，但共同特征是出生于 1930 年至 1935 年。

下图直观地表示了横向研究与 3 种纵向研究的基本逻辑：[1]

横　向　研　究

横向研究
1990 年
21～30
31～40
41～50
51～60

[1] 风笑天《社会学研究方法》，中国人民大学出版社，2001 年 3 月，第 83 页。

固定样本研究

	固定样本研究	
1990 年		2000 年
21 ~ 30 *	↖↘	21 ~ 30
31 ~ 40 *	↖↘	31 ~ 40 *
41 ~ 50 *	↖↘	41 ~ 50 *
51 ~ 60		51 ~ 60 *

趋 势 研 究

	趋势研究	
1990 年		2000 年
21 ~ 30	←→	21 ~ 30
31 ~ 40	←→	31 ~ 40
41 ~ 50	←→	41 ~ 50
51 ~ 60	←→	51 ~ 60

世 代 研 究

	世代研究	
1990 年		2000 年
21 ~ 30	↖↘	21 ~ 30
31 ~ 40	↖↘	31 ~ 40
41 ~ 50	↖↘	41 ~ 50
51 ~ 60		51 ~ 60

←→表示比较 * 表示同一组人



三、按数据收集方式分类

社会科学研究方式，按数据收集方式的不同可以分为定量研究与定性研究。定量研究侧重的是对研究对象的测量和计算，而定性研究则侧重于对事物的含义、特征、隐喻、象征的描述和理解。

在精确新闻中，我们所用的方法都属于定量研究的方法。有关定量研究和定性研究的差别比较如表2-3所示。有关定性研究的更详细内容请参见其他参考书籍。

表2-3　定量研究与定性研究的差别比较①

	定量研究	定性研究
哲学基础	实证主义	人文主义
研究范式	科学范式	自然范式
逻辑过程	演绎推理	归纳推理
理论模式	理论检验	理论建构
主要目标	确定相关和因果关系	深入理解社会现象
分析方法	统计分析	文字描述
主要方式	实验、调查	实地研究
资料收集	量表、问卷、结构观察等	参与观察、深访等
研究特征	客观	主观

定量研究可以分为实验研究、调查研究和内容分析。针对

① 风笑天《社会学研究方法》，中国人民大学出版社，2001年3月，第13页。

每个部分，在本书中会有专门的章节来介绍。这里只是介绍一下概要情况。

1. 实验

实验（experiment）方法是在精心加以控制的条件下，来研究两种变量之间的关系。按照实验地点的不同，实验可以分为实验室实验和实地实验。

在实验室实验中，研究对象和一切必要的材料均被置于一个被人为地安排好的，可以由研究者细心控制的环境之中。可以对情境的每个微小细节都加以控制。

实地实验则在实验室以外进行，条件比较自然，可能是在街道上进行，也可能是在医院或工厂里进行的。在精确新闻报道中，所采用的实验多是实地实验。

2. 调查

在社会学研究中，人们常常使用调查（survey）方法，目的是收集事实，或者是为了找出事实之间的关系。

要了解社会特点及其相互关系，调查是一个必不可少的信息源。不过，调查也有不足之处，如有时候，人们会对他们根本不了解的问题信口地发表意见；此外，回答率偏低等都会影响结果的准确性。

在精确新闻报道中，调查是最经常被采用的方式。据估计在美国每年有3000万人次接受调查访问，任何一户都接受过至少一次的访问[1]。我们会有专门的章节讲述调查方法以及电脑辅助的电话调查方法。

3. 内容分析

有时研究者不必或者不能通过实验或调查研究来获得新的

---

[1] Norman M. Bradburn & Seymour Sudman《Polls & Surveys-Understanding what they tell us》, Jossey-Bass Publishers, 1988年，第3页。



信息，只需要对现有的有关资料进行收集和分析。大量有用的信息都是可以获得的，有的是已经公开发表的，有的是尚未发表的，它们或者是由政府机关公布的统计资料、新闻短片、日记、信件、法庭记录、文艺作品，或者是其他社会科学家的研究成果。

内容分析（content analysis）是一种客观、系统地对文献资料进行整理分析的方法。

在社会学中利用现有资料进行研究的一个经典例子就是1897年迪尔凯姆（Emile Durkheim，又译涂尔干，1858～1917年）对自杀的研究。

迪尔凯姆想找出人们自杀的原因，他怀疑仅从个人心理方面来解释是不够的。在自杀问题上，显然不能进行实验，调查法当然也行不通。但是有关自杀的统计资料却唾手可得。

迪尔凯姆对这些数据资料进行了内容分析并加以统计，从而得出一些结论：自杀的发生率因社会群体而异，并且多年来一直如此；基督教徒比天主教徒更可能自杀；大城市里的人比小社区中的人更可能自杀；独自生活的人比在家庭中生活的人更可能自杀，等等。

上升到社会学的理论，迪尔凯姆从中分离出一个隐藏在这些差别后边的自变量：个人与他人实现社会结合的程度，与自己的社区关系淡漠的人比那些有比较密切联系的人更容易轻生。

后来，迪尔凯姆又对传统社会中一种比较常见的自杀形式进行了研究，也就是由于社会期待某个人自杀而出现的自杀。例如，在传统的日本，在某种情况下，自杀被视为一种高尚的行为；在传统的印度社会，社会有时候期待寡妇跳进其丈夫火化的柴堆，如果她这样做了，就会被看成是一个尽了自己应尽

责任和义务的妻子。在这些事例中，导致个人自杀的不是个人与社会的关系太弱，而是太强。这些社会里的人也许为了响应自己也赞成的社会期许才自杀的。

迪尔凯姆对自杀的分析，不仅使之成名，而且证明了社会学是一门科学。在精确新闻报道中，我们采用内容分析方法主要的目的是对所研究问题进行描述。

第二节　概念：定义及其操作化

一、概念、变量和指标

1．概念

概念（concept）是对现象的抽象，它是一类事物的属性在人们主观上的反映。比如"椅子"就是一个很简单的概念。但当说道椅子的时候，人们头脑中出现的特定的椅子是不同的：有的是木头的、有的是塑料的、有皮制的；有方的、有圆的；有棕色的、有黑色的，等等。尽管这些椅子不完全一样，但它们具有一个共同的特征：有若干条腿支撑着的、供人坐的、有靠背的一个平面。"椅子"这个概念正是对这些具体的、各不相同的椅子的抽象表述。

由于现实世界中事物和现象的类型、结构、复杂程度的不同，概念的抽象程度也是不同的。比如"责任心"、"生活方式"、"现代化"这样的概念，抽象的程度是高于"椅子"、"书"、"年龄"这类概念的。

在精确新闻中，编辑记者最感兴趣的概念主要包括信念、态度以及行为倾向等。信念是一个人对人对事的看法，如"北京是有能力举办2008年奥运会的"；态度是一个人对某些事件



或问题所持的正面或负面的感觉，如“我非常支持北京申办2008年奥运会”；行为倾向是指一个人是否有意采取某一项活动，如“我不会购买盗版影碟”。

2. 变量

变量（variable）也是一种概念。在社会科学研究中，我们经常把一个概念赋予两个或两个以上的数值，使概念在质或量有所变动。借助数学上的术语，把这样的概念称为变量，而把那些只有一个固定不变的值的概念，称为常量。变量不过是一种可以随着时空的变化而变化或者因人因群体而异的特点。年龄、性别、种族和社会阶级方面的差异都可以是变量。人们在智力、民族、收入和对某事的态度看法的差别也是变量。

变量具有两个重要的性质：

(1) 构成变量的各个值必须是穷尽的，即任何情况都可以归到某个取值中。比如文化程度这个变量，如果它的取值只有“初中”、“高中”、“大专以及以上”，则没有穷尽，没有包括初中以下文化程度；

(2) 构成变量的各个值必须是互斥的，即每种情况只能归到某一个取值中。比如，职业这个变量，如果取值中，有“工人”、“司机”、“车工”等，那么取值就不是互斥的了，而是相互包含的。

3. 指标

指标（indicator）是指可以直接测量的概念或者变量。比如，“文化程度”是一个概念，它又可以直接进行测量，所以，文化程度也是一个指标；而“社会阶级”是一个抽象的概念，不能直接进行测量，但可以通过职业、收入、文化程度等一组指标来间接测量。

表2-4是选自《旧金山编年史》（San Francisco Chronicle）

1982年1月13日的一篇报道，内容是关于在一场暴风雨中，两个地方所遭受的损失情况。

表2-4

	Marin Country	Santa Cruz
商家损失（万美元）	1500	5650
死亡人数	5	22
受伤人数	379	50
无家可归人数	370	400
房屋倒塌间数	28	135
房屋损坏间数	2900	300
商铺倒塌间数	25	10
商铺损坏间数	800	35
私人财产损失（万美元）	6510	5000
公共财产损失（万美元）	1500	5650

在上表中，所列的“死亡人数”、“受伤人数”等就是指标，共列出10项具体的指标，如果我们的研究目的是确定两个地区哪一个灾情最严重？那么答案应该是什么呢？

“灾情严重程度”是个抽象的概念，必须要用具体的指标来加以界定。如果“灾情严重程度”是指死亡人数，那么，Santa Cruz的受灾情况最严重；如果“灾情严重程度”是指受伤人数，那么，Marin Country的受灾情况最严重；定义“灾情严重程度”的指标不同，那么对“哪一个地区灾情最严重？”的回答是不同的。



二、概念的操作化

概念的操作化（operationalization）过程就是指把抽象的概念转化为可测量的具体指标的过程。操作化包括两个方面，界定概念以及建立测量指标。

1．界定概念

界定概念，就是对概念给出理论上的定义（theoretical definition）。由于我们的研究需要他人的理解和评判，我们必须对概念加以解释。所用到的概念越抽象，不同的人的理解可能差异越大，越需要给出定义。

举例来说，让人们写出对“年龄”这个概念的理解，大部分人写的都是“某人生存的时间长度”、“出生以来的经过的时间”以及类似的陈述。不同人之间对这个概念的定义重叠部分很高。

又比如，现在我们让人们来定义“媒介使用”，我们可能会得到不同的结果，有人认为是指使用时间，有人认为是指使用目的，等等，不同人之间对这个概念的定义重叠部分很少。“年龄”与“媒介使用”相比，要具体得多，而“媒介使用”是个比较抽象的概念，要根据具体的研究目的，给出具体的定义，如果研究关心的时间的分配问题，那么“媒介使用”可以定义成“使用媒介的时间”。

2．建立测量指标

概念的界定完成后，要给出具体的测量指标，使得此概念可以被测量。比如，“使用媒介的时间”可以细分为：看电视的时间、听广播的时间、读报纸的时间、看杂志的时间、看书的时间以及上网的时间。

有些时候，一些概念非常抽象，要从不同的方面或者说维

度来加以分解。我们来看一个例子：

1992 年在美国总统选举期间，美国 Connecticut 大学教授祝建华博士和 J. Ronald Milavsky 博士，利用总统候选人在电视上公开辩论的机会，进行了有关接触传媒对知识获取以及印象评价的影响效果的实验研究。

研究的内容包括对总统候选人的印象评价。那么，"印象评价"是一个抽象的概念，可从 5 个维度加以测量，即"能干（competence）"、"沟通技巧（communication skills）"、"正直（integrity）"、"个人魅力（charisma）"以及"权势（potency）"；这 5 个方面的评价还可以进一步细分，每个维度都从 3 个方面来评价的，具体如表 2-5 所示。

表 2-5

印象评价	能干	有才智	Intelligent
		能干的	Competent
		知识广博	He knows a lot
	沟通技巧	善于表达	Expressive
		条理清楚	Clear
		明确的	Speaks in specifics
	正直	值得信任	Trustworthy
		真诚的	Sincere
		诚实的	Honest
	个人魅力	举止文雅	Pleasant
		亲切的	Warm personality
		友好的	Friendly
	权势	敢作敢为	Aggressive
		强壮的	Strong
		支配的	Dominant



对于同一个概念，在不同的地点、不同的时间，不同的研究者所给出的操作性指标可能是不同的。

在家庭社会学中，“夫妻权力”是个常见的概念，研究对这一概念的界定是“夫妻在家庭中的决策权”。不同的研究者给出了不同的操作化指标。

表 2-6

[美国] 布拉德和沃尔夫 (1960 年)	[美国] 森特斯 (1971 年)
(1) 丈夫的职业选择	(1) 丈夫的职业选择
(2) 买什么样的汽车	(2) 买什么样的汽车
(3) 是否买人寿保险	(3) 是否买人寿保险
(4) 到什么地方度假	(4) 到什么地方度假
(5) 买什么样的房子	(5) 买什么样的房子
(6) 妻子是否应该参加社会工作	(6) 妻子是否应该参加社会工作
(7) 家里有人生病时，应去看哪位医生	(7) 家里有人生病时，应去看哪位医生
(8) 全家每周在食品方面应花多少钱	(8) 全家每周在食品方面应花多少钱
	(9) 请谁来做客和与谁一起出门
	(10) 怎样装饰房间和摆设家具
	(11) 收看/听什么电视/广播节目
	(12) 家庭的正餐吃什么
	(13) 买什么样的衣服
	(14) 配偶应买什么样的衣服

续表

[台湾] 陈明穗 (1986年)	[台湾] 伊庆春和蔡瑶玲 (1988年)
(1) 丈夫的职业选择	(1) 丈夫的职业或工作选择
(2) 妻子的职业选择	(2) 妻子是否外出工作或改变工作
(3) 家人生病该看哪个医生	(3) 家用支出分配
(4) 家庭生活费的支配	(4) 储蓄、投资、保险等
(5) 度假、旅游及休闲活动	(5) 婚丧喜庆时贺礼的数额
(6) 生育子女数目	(6) 买房子或房地产
(7) 购买贵重物品	(7) 要不要搬家
(8) 置产(买房子、土地等)	(8) 是否与上一代同住
(9) 房子布置及购买家具	(9) 生几个孩子
(10) 订阅报刊、选择电视节目等	(10) 用哪种避孕方法
(11) 子女的教育	(11) 小孩上哪所学校
(12) 谁代表参加婚丧喜庆	(12) 小孩教育
(13) 谁决定送礼与回赠礼物	(13) 请哪些客人吃饭
(14) 是否买保险	(14) 娱乐休闲活动的计划
(15) 何种家电用品该淘汰换新	(15) 家中的布置及买家具

在建立概念的测量指标时，应尽可能参考和利用前人已有的、比较成熟的、业界公认的指标。当然，也可以在前人指标的基础上，结合自己研究的目的，进行适当的补充和修改。如果，可供参考的资料有限，可以先进行探索性的研究，来收集资料，从而建立测量指标。



第三节　变量的类型及转换

按照不同的分类标准，变量可以分为不同的类型。这里我们介绍的是按变量之间的关系来分类的以及按照变量的取值的不同所区分的类型。

一、按变量之间的关系分类

按照变量之间的关系，可以把变量分为自变量、中间变量和因变量。

自变量（independent variable）是能够引起其他变量变化的变量，受到自变量影响的变量叫做因变量（dependent variable)。在美国的总统选举中，选民的财富多少是自变量，可以影响人们去投共和党候选人的票。

同一变量在某种情况下可以是自变量，在另一种情况下又可以是因变量。当驾驶的技术状况取决于酒醉的程度时，酒醉的程度就是自变量；而当酒醉的程度取决于饮酒量时，酒醉的程度又是一个因变量。

有时，自变量不是直接影响因变量的，而是通过中间变量(intervening variable）起作用的。如果饮酒量的多少看成是自变量，驾驶技术状况看成是因变量，则酒醉的程度就是中间变量。

自变量——→中间变量——→因变量
饮酒量——→酒醉的程度——→驾驶技术状况

二、按测量尺度不同分类

测量的尺度一般分为 4 个层次，从低到高分别是定类测量

(nominal measurement)、定序测量(ordinal measurement)、定距测量(interval measurement)以及定比测量(ratio measurement)。变量根据测量尺度的不同，可以分为定类变量、定序变量、定距变量和定比变量。

1．定类变量

定类变量本质上是一种分类体系，即把变量的不同属性或特征加以区分，标以不同的名称或符号，确定其类别。在社会科学研究中，对职业、性别、婚姻状况等特征的测量，所用变量多是定类变量。如性别分为“男”、“女”；婚姻状况分为“未婚”、“已婚”、“离婚”、“丧偶”、“其他”；

定类变量的数学特征是“等于”或“不等于”，或者说“属于”与“不属于”。

对应着定类变量的测量尺度叫做定类尺度。

2．定序变量

定序变量的取值是按照某种逻辑顺序把变量的取值排列出高低或大小，确定其等级次序。比如，文化程度变量，可以分为“初中及以下”、“高中”、“大专及以上”3个水平，这是从低到高来排列的；又比如，城市的规模，分为“特大城市”、“大城市”、“中等城市”和“小城市”，这是从大到小的顺序排列的。

在精确新闻报道中，所用到的变量以定序变量为主。如测量人们对某个事件的支持程度，多采用“非常支持”、“比较支持”、“无所谓”、“不太支持”、“非常不支持”的形式，也是一种定序变量。

对应着定序变量的测量尺度叫做定序尺度。

3．定距变量

定距变量不仅能够将社会现象或事物区分为不同的类别、



不同的等级，而且可以确定不同水平之间的间隔距离和数量差别。比如人的智商，测得张三为 125，李四为 130，我们就可以知道，李四比张三聪明；再比如温度这个变量，2001 年 5 月 20 日，北京最高气温为 37 摄氏度、最低气温为 21 摄氏度。我们可以知道北京的温差为 16 摄氏度。

定距变量的“零”值，不意味着“没有”，而只是一个给定的基准而已，或者说定距变量的零点是相对的。如温度为摄氏零度，并不是没有温度，而是温度刚好达到“冰水混合”的程度，也就是摄氏零度的规定标准。

对应着定距变量的测量尺度叫做定距尺度。

4．定比变量

定比变量除了具有定距变量的性质之外，还具有一个绝对的零点。所以定比变量不仅可以进行加减运算，还可以进行乘除运算。比如收入这个变量，直接询问具体的收入，就是一个定比变量，如果张三收入 1500 元，李四是 750 元，我们不仅可以说张三的收入比李四多 750 元，还可以说张三的收入是李四的两倍。对应着定比变量的测量尺度叫做定比尺度。

在社会科学研究中，多数情况下，我们不再区分定距变量和定比变量。

表 2-7　几种测量尺度的比较

特征	定类	定序	定距	定比
分类（=、≠）	√	√	√	√
次序（>、<）		√	√	√
距离（+、-）			√	√
比率（×、÷）				√

三、变量的转换

根据研究目的的不同，对于同样的问题，可以用不同测量级别的变量来测量。例如，"年龄"这个问题，可以采用以下几种方式询问：

1．"你的年龄是多少岁？"回答的是具体的年龄数字，属于定比变量；

2．"您的年龄属于以下哪一个年龄段？"

（1）21岁以下　　（2）21～30岁　　（3）31～40岁

（4）41～50岁　　（5）51～60岁　　（6）60岁以上

答案是在6个数字中选择，则为定序变量；

3．如果问："您的年龄是否大于18岁？"

（1）是　　（2）否

答案为"是"或者"否"，则为定类变量。

高级别的变量，可以转化为低级别的变量，反之则不行。比如说通过第一种问法，我们知道一个人的年龄是32岁，如果转化为第二种问法的话，答案为（3）31～40岁之间；如果转换为定类变量的话，则答案为（1）是大于18岁。反过来，我们知道一个人的年龄大于18岁，无法判断是在21～30岁之间还是51～60岁之间，当然，更无法转化为具体的年龄数字。

在实际应用中，要根据研究目的的不同，选择合适的测量级别。

第四节　量　表

在社会科学研究中，经常需要测量一些非常抽象的概念，如公众对某些问题的态度、看法、意见等主观性的内容，特别



是在精确新闻报道中，关心的多是人们的态度和看法。这些概念很难用单--的指标来测量，一般是采用量表的形式来进行复合测量的。

一、李克特量表

李克特量表（Likert scale），是在1932年由美国社会心理学家李克特提出来的。李克特量表也叫做总加量表，是由一组陈述组成的，这些陈述都表达了对某事的态度或看法，对于每种看法，有多个选项，来表示同意或支持的程度，常见的是分为“非常同意”、“比较同意”、“说不准”、“不太同意”以及“很不同意”5个级别，称为五级李克特量表。也有分成七级或者九级的。答案的级别越多，人们的态度差异就越能表现出来。

组成李克特量表的多项陈述中，每个陈述的强烈程度都是相同的。要求既有正向的陈述，也要有反向的陈述。所有陈述的得分相加，就构成了对某事的态度得分。

1990年第十一届亚运会在北京举行，对于亚运会的宣传效果进行了大规模的问卷调查，在问卷中，问及了人们对亚运会的态度，采用的是五级李克特量表。见表2-8：

表2-8

	说　法	选		项		
1	亚运会将显著地提高我国的国际声望，树立良好的国际形象	1	2	3	4	5
2	作为中国公民，我为亚运会在北京召开感到自豪	1	2	3	4	5
3	在国家财政十分困难的情况下不该举办亚运会	1	2	3	4	5

续表

	说　法	选		项		
4	只要能开好亚运会，即使影响我的生活、工作秩序也没关系	1	2	3	4	5
5	亚运会有利于振奋民族精神，激扬爱国主义	1	2	3	4	5
6	亚运会的召开能为我国带来可观的经济效益	1	2	3	4	5
7	花这么多钱办亚运会，还不如把钱用于改善人民生活	1	2	3	4	5
8	我国举办亚运会，充分显示了社会主义制度的优越性	1	2	3	4	5
9	亚运会的宣传声势虽大，但有些铺张浪费	1	2	3	4	5
10	我对亚运会后经济形势的发展感到忧虑	1	2	3	4	5
11	亚运会向老百姓集资，增加了人民的负担	1	2	3	4	5
12	亚运会将能展现我国安定团结和改革开放的重大成就	1	2	3	4	5

其中选项“1”代表“很不同意”，“2”代表“不同意”，“3”代表“说不准”，“4”代表“同意”，“5”代表“非常同意”。

很显然，在有关亚运会态度的12项陈述中，第1、2、4、5、6、8、12条陈述为正向的，而第3、7、9、10、11条陈述为反向的。在计算对亚运会的态度得分时，应该先把反向陈述的得分转换一下，即“很不同意”得5分，“不同意”得4分，“说不准”仍然得3分，“同意”得2分，而“非常同意”得1分。然后把对12项陈述的得分相加，就构成了对亚运会的态度得分。



二、鲍格达斯社会距离量表

鲍格达斯社会距离量表（Bogardus social distance scale）是由在内容上具有某种趋强的逻辑结构的一组陈述构成的，与总加量表不同的是，每一条的陈述代表着人们在态度上的强烈程度是不同的。

下表测量的是人们对黑人的态度。不同的问题所显示的人们相互之间的距离的不同，越往后，表示人们之间的距离越近。或者说，同意后面的陈述的人，一定会同意前面的陈述。一个愿意让子女同黑人结婚的人，一定会同意前面的5条。但是，一个愿意与黑人做邻居的人，却不一定愿意与黑人交朋友。见表2-9：

表 2-9

	愿意	不愿意
1．你愿意让黑人生活在你的国家吗？	□	□
2．你愿意让黑人生活在你所在的城市吗？	□	□
3．你愿意让黑人住在你所在的街区吗？	□	□
4．你愿意让黑人做你的邻居吗？	□	□
5．你愿意与黑人交朋友吗？	□	□
6．你愿意让你的子女和黑人结婚吗？	□	□

用鲍格达斯社会距离量表可以测得不同的人或群体对某一群体的态度。

三、语义差异量表

语义差异量表（semantic differential scale）的形式由处于两

端的意义相反的词（多是形容词）构成，每一对反意词中间可分为7个等级或者是9个等级。语义差异量表主要用于文化的比较、个人及群体间差异的比较以及人们对事物的态度、看法等方面的研究。

1992年美国总统选举期间，祝建华等人选用语义差异量表来评价对总统候选人的印象。对每一位总统候选人，用15对反意词来加以形容。见表2-10：

表2-10

中文意义	“7”的含义								“1”的含义
有才智	Intelligent	7	6	5	4	3	2	1	Not Intelligent
能干的	Competent	7	6	5	4	3	2	1	Incompetent
知识广博	He knows a lot	7	6	5	4	3	2	1	He doesn't know much
善于表达	Expressive	7	6	5	4	3	2	1	Unexpressive
条理清楚	Clear	7	6	5	4	3	2	1	Vague
明确的	Speaks in specifics	7	6	5	4	3	2	1	Speaks in generalities
值得信任	Trustworthy	7	6	5	4	3	2	1	Untrustworthy
真诚的	Sincere	7	6	5	4	3	2	1	Insincere
举止文雅	Pleasant	7	6	5	4	3	2	1	Unpleasant
亲切的	Warm personality	7	6	5	4	3	2	1	Cold personality
敢做敢为	Aggressive	7	6	5	4	3	2	1	Passive
强壮的	Strong	7	6	5	4	3	2	1	Weak
支配的	Dominant	7	6	5	4	3	2	1	Subservient
友好的	Friendly	7	6	5	4	3	2	1	Unfriendly
诚实的	Honest	7	6	5	4	3	2	1	Dishonest



第五节 测量的信度和效度

信度（reliability）表示测量的可靠性（trustworthiness），是指测量结果的稳定性（stability）和一致性（consistencies）。

测量的效度（validity），也叫做测量的有效度，是指测量工具所测量的概念正是所要研究的概念而不是其他的概念，并且这些概念被准确地进行了测量。或者说，效度是指实际的测量在多大程度上反映了概念的真实含义[①]。

对一个人的体重进行测量，用同一个地秤，如果1分钟内测量两次，这两次的结果如果有很大的差异，则说明这个地秤的信度低；对一个人的身高进行测量，如果测量多次结果相同，则说明信度高，但如果这个人是穿着鞋量身高的话，则效度低，即所测得的高度不仅包括身高还包括鞋底的厚度。

在我们前面所举的对亚运会的态度测量中，研究的目的是了解人们对亚运会的态度，这个态度得分是客观存在的，但未知的。我们是采用了一个包括12项陈述的五级李克特量表来测量的，如果所测得态度得分就是真实的态度分，我们就说此量表是有信度的。

我们的目的是了解人们对亚运会的态度，但所测得态度可能既包括对亚运会的态度，也包括对其他事情的态度。如果通过量表所测得态度都是关于亚运会的态度，那么我们就说这个量表是有效度的。

不仅对于量表有信度、效度的问题存在，对于整个研究设计，对于具体的数据收集方法，包括调查、实验和内容分析，

① ［美］艾尔·巴比《社会研究方法》，华夏出版社，2000年4月，第172页。

都可以从信度和效度两个方面来加以评价。但对于不同的内容，信度、效度的评价侧重点是不同的。比如，对于实验方法，更多是从内部效度、外部效度两个方面加以评价，而对于内容分析，多是从编码员的信度来评价的。

一、信度

在社会科学中，要确定测量是否可信是一件困难的事。因为要是将原测量工具，比如问卷，对研究对象再次使用，研究者本是想以此来核对测量手段的信度，但研究对象可能会认为研究者对他们过去的回答不相信，要检验一下他们是否撒谎。因此，被访者可能有意地做出无矛盾的回答。这种情况会破坏对一种测量手段和方法的重复使用。

目前对信度检验的方法，一般采取以下两种方式：

1、交替或平行法

交替或平行法类似于用两个一样的秤在同一时间内测量同一个人的体重，从而相互印证它们的可靠性。

例如：研究者可以设计出测量同一概念但具体问题不同的问卷，同时用于研究对象，从结果上相互印证它们的可靠性。这种方法又叫做多形式法。

另一种方法叫折半信度法。研究者使用此法只需设计出单一的测量问题，它包纳的问题是研究者实际需要的两倍，其中的一半是对另一半的重复。这两半都同时在研究中使用，其结果如果高度相关，便可以说该测量手段是可信的；如果相反，便可以说该测量手段不可靠。

使用上述方法是要弄清楚设计的两种交替手段或两个对半测量条目确实测量的同一概念。如果不是这样，即它们各自测量了不同的概念，研究者便不能对信度做出评估。就实际意义



来讲，使用两种相互替代的测量手段，只能估计出它们是否测量了同一概念的程度。如果这两种测量结果是高度相关的，便可以说明都测量了同一概念；如果相关是不显著的，则说明测量了不现的概念。也有人认为，用交替方法并非是对信度的检验而是对准则效度的检验[①]。

2. 信度的反复检验

由于对信度的规定指的是同一测量尺度其结果的一致或者无矛盾，因而检验信度最好是对同一测量手段的方法反复运用，即测量，再测量（Test-retest）。但这样做会碰到许多问题。

比如在日常生活中，一个人如果使用自己家里的地秤每日称一次体重，连续两个月，也许会发现最后的记录比第一天的重了3公斤，这不一定是测量不可信。因为这个人可能在两个月体重增加了。因而精确的地秤便显示超出原有的体重。

一个有信度的测量手段，会在被测对象的特征发生变化时，显示其特征值的变化。如果没有变化，就不会做相应的反映。

有些初接触研究方法的人认为，一种测量手段要成为可信的，必须在各种使用的时期都表现出其测量值的一致。这只是被测量对象无实际变化时才对。而有的测量手段即使始终显示其测量值的一致，也不一定可信。如一支坏了的温度计，无论冬夏，都显示一个数值，人们也不会相信的。

在社会科学研究中，这个问题在于弄清：在什么时候重复使用同一手段进行测量应允许出现不一致的结果；所得到得测量值是反映了真实的变化还是测量手段不可信。

---

① 袁亚愚、徐晓禾编译《当代社会学的研究方法》，四川大学出版社，1986年9月，第38页。

三、效度

对测量来讲，常用的效度包括表面效度、准则效度和结构效度。

1. 表面效度

表面效度（face validity），也称内容效度或逻辑效度，它是由评估人对测量的效度做出的一种直接的判断。反映的是专家的意见。

评估人一般是研究领域的专家，评估人对被测量的概念进行研究，以决定所用的测量手段能否对研究概念进行有效测量。一个测量是否具有表面效度可以通过专家来判断，但不能被证明。

要判断一种测量手段是否具有表面效度，首先要了解被测量概念的定义，其次要弄清楚收集的资料是否与被测概念紧密相关。在如果测量手段所测量的正是设想中的概念而不是其他概念，该测量手段具有表面效度；否则，就无表面效度。

如果对被测量的概念定义不清楚，或者测量过程长并且复杂，就很容易产生表面效度问题。

2. 准则效度

准则效度是运用一个已有的、公认的测量尺度作为检验新的测量尺度是否有效度的标准。在进行新的测量中，如果新尺度测量的结果与作为标准的旧测量尺度的测量结果是相同的或类似的，就可以说新的测量尺度有准则效度。

要注意的是，作为衡量标准的旧的测量尺度，必须是具有表面效度，并且已在实践中被证明是有效的。

既然旧的测量尺度是有效的，为什么不采用它而要用新的呢？具体的原因很多，如旧尺度实用困难、实施困难，以及难



以编码，甚至用语陈旧、过时等等。由于不同的原因，使得旧的测量尺度不再适应新情况的需要，所以，便提出了设计新的测量手段的必要。

3．结构效度

结构效度是由两个相关的可以相互取代的测量尺度，对一定理论命题交互测量可取得同样的测量结果而构成的。当运用第一种尺度对理论、命题和概念进行测量而都被事实证明后，又用第二种尺度去代替第一种并取得同样结果，就可以说新的测量尺度（第一种）有结构效度。

举例来说，在社会学中，我们要测量人们对婚姻的满意程度，假设我们已知对婚姻的满意程度与主动做家务的行为密切相关，对婚姻满意程度越高的人，越是主动做家务。如果我们的测量也表明婚姻满意程度与做家务密切相关，那么，对婚姻满意的测量就具有结构效度；如果测量发现，对婚姻满意程度不同的人在做家务行为上没有差异，则测量的结构效度存在问题。

从表面效度到准则效度再到结构效度，可以被看成是递进的上升加过程。结构效度要求比准则效度、准则效度要求比表面效度有更多的情报。因此，结构效度常常被认为是各种效度鉴别中最高的一类。

在实际的鉴别中，表面效度只涉及单个的概念和关于该概念的单个测量尺度，而准则效度则需一个概念和测量它的两个或更多的测量尺度。至于结构效度，不仅要涉及某一概念及其相对的两种以上的尺度，而且还要设计另外的概念及它们的测量尺度。这些尺度是能够通过命题与问题中的概念相联系的。

三、信度和效度的关系

测量的信度和效度既相互影响，又相互制约。

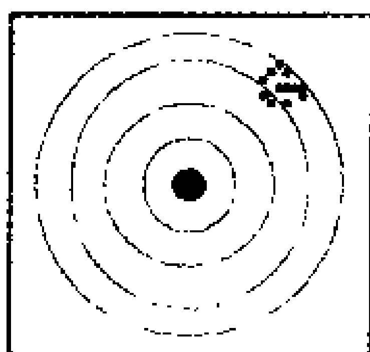
信度高，效度却不一定高。

如在对学生的身高的测量中，量身高时学生们都是穿着鞋量的，测量几次，身高都是一样的，很显然，信度高，但效度不高。

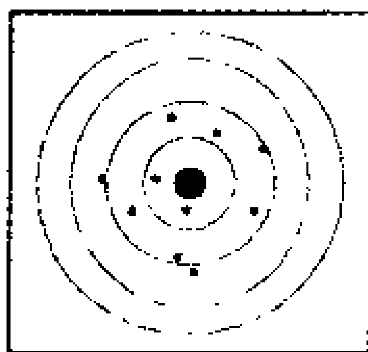
效度高，但信度不一定高。

在社会学中，测量家长对孩子的溺爱行为，可以用问卷调查的方法，也可以采用实地观察的方法。采用调查的方法，信度会高，但效度可能比较低。如果采用实地观察的方法，深入到每个家庭中仔细观察，这样做的效度比较高，研究人员实实在在地看到和感受到家长们是如何培养孩子的，但信度可能比较低。[1]

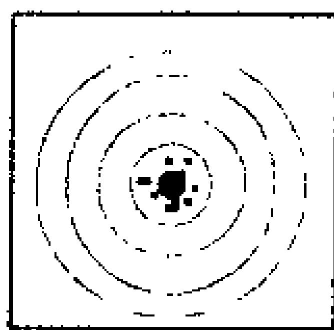
下面3个图直观地表示了信度和效度的关系。



信度高，效度低



效度高，信度低



信度高，效度高

① 风笑天《社会学研究方法》，中国人民大学出版社，2001年3月，第112页。



第三章　抽　样

抽样是指从我们所要研究对象的全体中，抽取一部分对象进行调查。如何抽取这一部分对象，是十分重要的，关系到从这一部分对象中得到的结论能否推广到所研究的全部对象。在精确新闻中，常用到的有关总统选举的预测调查，就是只对一部分选民进行调查，通过这一部分选民的投票意向，来预测在大选中，到底谁会当选？要预测得准确，就要求所调查的那部分选民的投票意向对全体选民的投票意向有代表性。

1936年美国总统大选，两位候选人分别是罗斯福（Franklin Roosevelt）和兰登（Alf Landon）。大选前，《文学文摘》（Literary Digest）杂志按照电话号码簿以及汽车牌照登记簿上的地址，寄出了1000万张模拟选票，共回收了230多万张模拟选票，结果显示，57%的人将投票给兰登，只有43%的选民会支持罗斯福。于是，《文学文摘》根据此次调查结果，预测兰登会在大选中击败罗斯福，当选美国总统。

同时，盖洛普采用了配额抽样的方法，只调查了3000人，预测罗斯福将会在大选中以55.7%的选票获胜。盖洛普考虑了地区、性别、年龄、所支持的政党以及收入等多种因素，按照这些因素来确定不同类型的选民各调查多少人。

选举结果我们已经知道了，在大选中，罗斯福获得了62.5%的选票，连任美国总统。《文学文摘》由于预测错误，声名受损，不久就被迫停刊了；盖洛普从此名声大振。

1936年的对总统选举的民意调查活动，不仅促进了民意调查的发展，同时也使得科学的抽样方法受到重视。

《文学文摘》的错误在哪里呢？除了邮寄方式以外，样本的选取是错误的根源。1936年，在美国社会中，拥有住宅电话和汽车的人数比例都很低，限于社会的中上层，而受1933年开始的美国经济大萧条的影响，大量的人口滑落到了中下层，因此，《文学文摘》所抽选的受访对象，绝大多数是中上层民众，遗漏了社会的中下层，而社会的各个阶层都是有投票权的。而且当时的中上阶层比较富裕、保守，反对罗斯福民主党政策的人比较多；而中下层希望选个民主党人当总统，投票的比例很高，大选中，这些没有被抽选进入样本的收入较低的中下层选民压倒多数地把选票投给了罗斯福，使《文学文摘》的预测失败。也就是说，《文学文摘》所抽选的调查对象对全体选民是没有代表性的。

盖洛普的配额抽样方法，在我们现在看来，也不是最科学的方法，但比起《文学文摘》的抽样方法，代表性要强很多，所以能正确预测罗斯福当选。但是预测的罗斯福获得的选票比例与实际的比例差距还是比较大的。

随着科学的抽样方法的发展，民意调查所采用的抽样方法也越来越科学，对总统选举预测的正确性也越来越高。

1984年11月15日美国总统大选的前一天，盖洛普民意测验根据对1985位可能投票的选民的调查结果，预测美国当时的总统，共和党候选人里根将会在大选中，获得59%的选票，击败民主党候选人蒙代尔，连任美国总统。11月16日大选结



果揭晓，里根获选票 53,428,357 张，占所有选票的 59.1%；蒙代尔获选票 36,930,923 张，占 40.9%。预测结果和实际结果相差只有千分之一。

在 1992 年的美国总统选举中，多家机构在大选前做过选举预测，预测结果与实际的选举结果都是吻合的。[①] 这一切都得益于科学的抽样方法。见表 3-1：

表 3-1

百分比%	克林顿（Clinton）	布什（Bush）	皮尔特（Perot）
最终大选结果	43	38	19
USA Today/CNN/Gallup	49	37	14
The Harris Poll	44	38	17
ABC News	44	37	16
NBC News/Wall Street Journal	44	36	15
CBS/ New York Times	44	35	15
Washington Post	43	35	16

本章将介绍抽样方法，包括抽样调查原理、抽样中常用到的一些基本概念，常见的概率抽样与非概率抽样的组织形式。在应用实例中，介绍了一个实际的抽样方案的设计。精确新闻中，最常用的调查方式是电话调查，有关电话调查中的抽样方法及应用实例，我们放在了第七章电脑辅助的电话调查 CATI。

① Roger D. Wimmer & Joseph R. Dominick《Mass Media Research-An Introduction》Wadsworth Publishing Company 第五版，1997 年，第 62 页。

第一节　抽样调查的基本原理

在设计一个研究的时候，必须先考虑很多问题，这项研究是采用抽样调查的方式还是普查的方式？是否有合适的抽样框使我们能接触到感兴趣的总体？抽样单位是什么？可能的误差来源是什么？误差对调查结果可能会造成什么样的影响？等等，这些问题我们必须事先考虑，一项调查或内容分析是否可行将取决于对以上所述问题的回答。

一、普查与抽样调查

在抽样设计中，首先要确定是采用普查的方式还是抽样调查的方式。

普查（census）也称全面调查，是指调查总体中的所有单元。

抽样调查（survey sampling）只是调查总体中的一部分单元，经常是很小的一部分。抽样调查也称样本调查（sample survey)，是非全面调查中的一种重要的方法，是按照一定的程序从所研究的对象的全体中抽取一部分进行调查或观察，获得数据，有时还要据此对总体的一定目标量进行推断。

一项研究，到底是采用抽样调查方式还是普查的方式，取决于许多因素，包括研究经费的多少、可动用的人力物力的多少、要求的时限以及总体的大小等等。在实际工作中，要根据具体的情况而定。

具体到精确新闻中所涉及的调查，绝大多数时候是以人作为调查对象的，如果总体比较小的话，比如，国外常做的对议会议员进行调查，由于议员的数目不大，就可以采用普查的方



式，即调查访问每一个议员；但是如果总体十分庞大，比如说是选民。在美国，选民的总数大概在1亿人左右，那么采用普查的方式的话，操作起来十分困难；另外，精确新闻既然是新闻的一种报道形式，同所有的新闻一样，对时效性的要求是比较高的，在最短的时间内获得调查结果是十分必要的。所以，在这种情况下，就应当采用抽样调查的方式。

二、为什么经常采用抽样调查

在日常生活中，人们自觉或不自觉地应用抽样方法，比如说，在厨房里炒菜、煲汤，起锅前总要尝一尝味道如何、咸淡如何；去市场上买花生、瓜子，也总是要抓几颗尝尝。在生产企业，对产品也要进行抽样，来检定产品质量；商家在进货时，也会抽查一部分，来判断是否合乎合同要求等等。

近几十年来，随着抽样调查理论和实践的发展，抽样调查的应用范围越来越广泛。在人口调查、经济调查、社会调查、卫生调查以及卫生环境调查领域中被普遍应用。

与普查相比，抽样调查有那些优点呢？或者说为什么经常采用抽样调查的方式？

1．有些总体是无限的

有些总体是无限的，如大气、水等，无法进行普查。比如说，在空气质量的预报中，我们要研究大气中可吸人颗粒物是多少，我们周围的大气是无限，只能采取抽样的方法；另外，如果我们想了解河流的污染情况，不可能对河流中的所有水进行检测，只能在有代表性的河段，抽取一定数量的水的样品，来进行检测化验。

2．总体很大

在社会经济研究和民意研究中，总体虽然是有限的，但过

于庞大，理论上可以进行普查，但实际上操作起来十分困难，甚至是不可行的。

3. 有些调查/实验是破坏性的

这种情况多发生在工农业生产中，比如说，我们要了解一批种子的发芽率，只能是从中抽取一小部分种子种下去；另外，工业生产中的产品质量、产品寿命的检验，如电视显像管的寿命的测量，是把一部分显像管处于工作状态，直到出故障或报废，测量才算结束。很显然，这样的破坏性检测只能对产品的很小部分来实施。

4. 抽样调查更经济

与普查相比，抽样调查节省人力、物力和财力。以我国1983年进行的第三次全国人口普查为例，正式参加普查工作的人员总计有518万名普查员，109万名普查指导员，13万名编码员，4000多名电脑录入人员，1000多名电脑工作人员，并得到了1000多万基层干部群众的配合与协助。在不包括普查工作人员的工资和劳务费的情况下，共花费人民币4亿元左右，以及联合国资助的1560万美元。

在抽样调查中，由于只对总体的一部分，通常是很少的一部分进行调查。这就使得抽样调查所需要的人力、物力大为减少。

5. 抽样调查可迅速获得结果

普查由于涉及对象多，组织工作复杂，调查的周期一般比较长。以我国1983年第三次人口普查为例，从1979年底国务院人口普查领导小组成立起，到1985年11月公布调查结果，宣布普查工作正式结束止，历时近6年。而抽样调查涉及的对象少，工作量小，周期短，时效性强。

6. 抽样调查的准确性可能更高

调查结果的准确性是指其与真实结果的一致性程度。理论



上说，普查调查了总体中的所有个体的情况，所以对总体状况的反映是最准确的；而抽样调查只调查总体中的一部分个体，对总体的反映程度不及普查，这部分误差统计上叫做抽样误差(sampling error)。也就是说，普查是没有抽样误差的，而抽样调查是存在抽样误差的。

但是在实际调查中，由于存在各种各样的人为误差，使得调查的质量下降。这部分人为的误差，统计上叫做非抽样误差(Non-sampling error)。普查由于涉及的时间长、涉及的人员多，所以非抽样误差可能更大。而抽样调查中非抽样误差可能很小。

误差越大，准确性越低，误差是由抽样误差和非抽样误差两部分组成的，普查没有抽样误差，但非抽样误差很大；抽样调查有抽样误差和非抽样误差，但非抽样误差很小，所以，一般说来，抽样调查的误差可能低于普查，即抽样调查的准确性要高于普查。

三、概率抽样的基本原理简述

在实际工作中，虽然多数时候我们采用的是抽样调查的方式，但是，研究的目的都是利用收集得到的样本数据的信息，对总体的情况进行估计。比如说，在国外的总统选举的预测调查中，只调查了几千人的投票意向，但研究者所关心的并不是被调查的人的意向，而是他们所代表的所有选民的投票意向。

我们所研究的总体是由不同的个体所构成的群体，如果总体中的每个个体在所有方面都完全相同，也就是说这个总体具有百分之百的同质性（homogeneity），这时，只了解一个个体，就可以了解整个总体的情况。但很显然，这是一种十分理想的情形，现实中绝大多数群体并不具备这种特征。现实中，总体

中的个体总是具有这样或那样的差异，即存在着不同程度的异质性（heterogeneity）。

研究总体很大，抽样调查中只抽取一部分个体来调查，这些个体的特征必须和总体的特征越接近越好，即样本要对总体有代表性。概率抽样的方法可以帮助我们实现这一目标，并对样本对总体的代表性程度做出估计。概率抽样的关键是随机抽取（random selection），就是要保证总体中的每一个个体都有相同机会被抽中人选样本。并且，总体中的各个个体是相互独立的，即某一个体是否被抽中，不受其他个体的影响。

概率抽样之所以能保证样本对总体的代表性，其原理在于概率抽样能很好地按总体内在结构中所蕴含的各种随机事件的概率来构成样本，使样本在一定程度上成为总体的缩影。

在统计学教材中，经常采用蒙特卡罗（反复模拟）的方法来归纳抽样分布，即随着样本量的增大（通常大于30），无论总体的分布如何，样本的均值、比例越来越接近于服从正态分布，这条定理在统计上叫做正态近似定理，在数学上称之为中心极限定理。详细的证明与解释请参看概率统计方面的教材。

必须注意的是，只有采用了概率抽样的方法，所抽取的样本才对总体有一定的代表性，根据样本所得到的结果，才能去推断总体的情况。如果是非概率抽样的话，用样本的结果去推断总体的情况，则风险很大。

四、概率抽样方案设计应遵循的原则

著名的美国抽样专家科什（Kish）教授在其代表作《抽样调查》（中国统计出版社，1997年出版中译本）中提出进行概率抽样设计时应该遵循的4条原则，即目的性原则、可测性原则、可行性原则以及经济性原则。



1．目的性原则

目的性原则是指在进行抽样方案设计时，要以研究的目标为依据，以研究问题为出发点，从最有利于研究资料的获取、最符合研究目的等因素来考虑抽样方案的设计。

2．可测性原则

可测性原则指的是抽样设计能够从样本自身计算得出有效的估计值或者抽样变动的近似值，通常以标准误差来表示。这是统计推断的基础，是样本结果与未知的总体值之间客观、科学的桥梁。通常，只有按照概率抽样的原则获取的样本才可以计算出有效的估计值。但是，概率抽样也不是自动保证可测性。比如，从一个具有周期性变化的总体中选出一个系统样本，就不能保证这种可测性。

3．可行性原则

可行性原则是指所设计的抽样方案必须在实际操作中切实可行。可行性意味着研究者所设计的方案能够预料到实际抽样过程可能遇到的各种问题，并给出相应的解决办法。理论上设计再好的抽样方案如果在实际操作中不可行，也是没有用的。

4．经济性原则

经济性原则是指抽样方案的设计要与研究所能得到的资源，包括费用、时间和人力等，相适应。

这4条原则是相互制约的，在实际工作中，要根据具体的情况进行取舍和平衡。

五、样本量的确定

样本量（Sample size）指样本中所包含的个体单位的数量。

1．决定样本量的因素

一般说来，决定样本量的大小要考虑下面几个因素：

（1）研究目的：如果一些研究是探索性或预测性的，则样本量不必太大。

（2）费用成本：费用的多少决定了样本量的多少。一般说来，样本量越大，费用越高。

（3）时间周期：一般说来，样本量越大，数据收集的时间周期就越长。

（4）总体的同质性：总体的同质性越高，抽样误差比较小，样本量可以小一些。

（5）子总体推断：如果分析时需要把被访者按某个标准区分为不同的组的话，则建议人数最少的组至少应该包括 50～100 人才算合适。[①]

2．非常简单随机抽样所需的最小样本含量

用于精确新闻的民意调查，总体一般很大，而样本相对较小。比如，在美国的总统选举调查中，总体有 1 亿人左右，而样本不过是 3000 左右；在收视率调查中，以北京为例，大概有 400 万户家庭，收视率的调查样本只有 500 户左右。所以，在民意调查中，有放回的非常简单随机抽样与无放回的简单随机抽样可以视为等同。

有专门的公式可以计算所需的样本量的大小。根据研究的目的，要先给出置信度水平（标示可靠程度）和最大允许的误差水平。在社会科学研究中，置信度经常取 90%、95% 和 99%3 种水平。只要将确定下来的置信度、最大允许的绝对误差带入公式，就可以得到简单随机样本所需要的最小样本量。

表 3-2 给出在置信度为 90%、95% 和 99% 的时候，最大允许绝对误差为 1% 到 7% 的时候，非常简单随机抽样所需要的

① 罗国辉《精确新闻报道》，第 101 页。



最小样本含量。第一行为置信度水平，第一列为最大允许绝对误差。

表 3-2

样本量	90%	95%	99%
1%	6806	9604	16641
2%	1702	2401	4160
3%	756	1067	1849
4%	425	600	1040
5%	272	384	666
6%	189	267	462
7%	139	196	340

比如说，中央电视台 1990 年筹建“中央电视台观众调查网”时，置信度要求是 95%，误差要求不超过 3%。[①] 参照上表，所需的样本量为 1067。

3．设计效应

要注意的是，上表给出的是非常简单随机抽样所需要的最小样本量。在实际工作中，大规模的民意调查所采用抽样方式都不是简单随机抽样，而是多阶段的抽样，所以抽样误差要大一些，为了保持所要求的精度，就要增加样本量。

设计效应表示的是具体的抽样方式（如整群抽样、多阶段抽样）与相同样本量的简单随机抽样的方差比，是事后才能计算出来的。在实际工作中，在事前的样本量设计中，根据经

① 柯惠新等《调查研究中的统计分析法》，北京广播学院出版社，1992 年，第 306 页。

验，设计效应一般取2、2.5、3这样的数。

对于多阶段抽样的民意调查，所需要的有效样本量可用以下公式计算：

多阶段抽样的样本量＝设计效应×非常简单随机抽样的样本量

实际应用中，利用以上公式计算出来的样本量一般作为有效样本量，即访问成功的样本量要达到这个数目。在美国三大广播网所进行的民意调查中，样本量一般在1200～1500之间[①]。

在实际访问过程中，有拒绝接受访问的情况存在，如果是出于确定费用和人力的预算而计算样本量时，还要考虑到访问成功率的大小。访问的成功率要根据以往的经验进行估计。不同的地区、不同的时间、不同的数据收集方式都可能会影响到成功率的大小。

多阶段抽样的样本量＝设计效应×简单随机抽样下的样本量÷访问成功率

第二节　抽样调查的基本概念

为了更好地理解抽样，我们先来介绍一下抽样中所要用到的一些概念。

一、总体、个体、样本

1．总体、个体、样本的概念

总体（population）是所研究对象的全体，由多个个体

① Norman M. Bradburn & Seymour Sudman《Polls & Surveys-Understanding what they tell us》，Jossey-Bass Publishers，1988年，第127页。



（individual item，也称个体单位）组成。样本（sample）是总体的一部分，是由从总体中按一定程序抽取得部分个体构成的，样本也是个体的集合。

总体可能是一群人、一批文件、一些事件等等；总体是个相对的概念，研究目的不同，总体的大小不同。总体的限定是人为的，是根据研究目地而定的。

新闻媒体进行民意调查时，主要的目的是了解广大民众对某些人物或事件的意见以及态度，因此个体通常是人，所抽取接受调查访问的那部分人就构成了样本。

在国外进行的总统选举预测中，总体就是所有的选民，从数量上说，可能有上亿；个体就是选民，所调查的选民就构成了样本。

如果研究的是北京近两年来造成交通事故的原因，则总体就是交通管理部门的近两年中的所有有关交通事故的记录，每一条记录就是一个个体。

如果想了解《读者》杂志 20 年来的内容，则总体是由《读者》杂志 20 年来的每一期刊物构成的。

如果想了解北京广播学院学生对食堂的意见，则总体是全体广院的学生，可能有几千人；如果想了解新闻系 98 级学生对某门课程讲课教师的意见，则 98 级新闻的全体同学就是研究总体，只有 40 人。

如果想了解《北京青年报》的精确新闻的报道风格，则首先要界定什么样的报道算做精确新闻，研究总体就是《北京青年报》的全部精确新闻报道，每一篇精确新闻报道就是一个个体。

对于任何一项调查，调查对象必须是明确的，也就是说，要对研究总体以及个体给出明确的界定。



2. 总体的类别

总体又可分为目标总体和被抽样总体。

目标总体就是期望从中获得信息的总体，它是我们要研究的个体单位的集合。目标总体可以是概念性的，并不需要一个实际存在的目录。比如说“所有在北京生活超过5年的人”、“年营业额在500万元以上的制造业企业”。

被抽样总体是实际调查所覆盖的总体，它可能与目标总体不是完全一致，但理想情况下，这两个总体应该是非常相近的。

目标总体和被抽样总体之间的差别源于多种的原因。比如，目标总体中的部分个体由于调查成本太高、调查实施起来太困难，不能成为被抽样总体的一部分。

需要着重说明的是，因为目标总体与被调查总体之间存在着差异，所以，依据调查结果所得出的结论，仅适用于被抽样总体，而不适用于目标总体。所以应该尽量缩小两个总体的差异。

二、抽样单元

总体是由个体组成的，总体可以是有限的，即由有限的个体组成，总体也可以是无限的，即总体中所包含的个体数量是无限多个。从抽样理论上讲，我们总可以把总体看成是有限的，可以将总体划分为互不重叠、又穷尽的有限多个部分。每个这样的部分称为一个抽样单元（sampling unit），也称为抽样单位。每个抽样单元是由或多或少的个体组成的，当然也可以只包含一个个体。

将总体划分为抽样单元是为了方便抽样。抽样单元可大可小。事实上，抽样单元可以分级。总体先分成若干个较大规模的抽样单元，称为初级抽样单元；每个初级抽样单元又可以包



含若干个规模较小的单元，称为次级单元，也称二级单元；以此类推，可以定义三级单元、四级单元等。最小一级的抽样单元就是个体单位，也称为基本抽样单元。

抽样单元可以是人为划分的，也可以是自然形成的。在社会调查中，各级行政单位常被作为各级抽样单元。

例如，在一项全国居民环境意识的调查中，可以将省（自治区、直辖市）作为一级抽样单元，将省内的市县区作为二级抽样单元；将街道、乡、镇作为三级抽样单元；将居委会或村委会作为四级单元；五级抽样单元就是居民家庭户；而更小的抽样单元就是个人了。



三、总体参数与样本统计量

每项调查都有其特定的内容和目的，通常调查的目标量都是用总体的某些指标来表示的，这些指标虽然是我们的目标量，是未知的，但都是客观存在的，所以有关总体的指标我们一般叫做总体参数。

最常见的总体参数有以下几种类型：

1. 总体总量

总体总量，又称总体总和（population total），比如说，我们调查的目的是想了解全国有多少网民，那么网民的数量就是一个总量。中国互联网络信息中心第七次全国调查报告（2001年1月）估计中国的网民数为2500万人。

2. 总体均值

总体均值，又称总体平均数（population mean），如职工的平均收入是多少，每月的伙食费支出平均是多少，考试的平均成绩，男性和女性的平均身高等等，都属于均值。

3. 比例

总体具有某种特征的个体在全体中的比例（proportion），常用百分率或成数表示。北京市的居民中，支持北京申办2008年奥运会的比例是多少？在精确新闻的民意调查中，比例问题出现的最多。

为了估计总体参数的大小，如果我们采用了概率抽样调查的方式，则根据所抽取的样本可以计算相应的指标，这些样本的指标我们叫做样本统计量。

四、抽样框

研究总体的定义明确以后，我们需要接触到总体的个体，抽样框（sampling frame）就提供了一个途径，来确认和接触被抽样总体中的单位。

另外，抽样框还可以告诉我们总体中个体单位的数目，这样的话，样本量确定以后，就可以计算出总体每个单位被抽中的概率，并根据抽样调查的结果对总体进行推断。

抽样框可以是一份名单，也可能是地图等资料。比如说，对北京广播学院的教师住房情况进行调查，那么，人事部门的



教师名录就是抽样框。原则上，抽样框的抽样单元不应该重复，也不能遗漏。北京广播学院教师名录中要包括学院所有的教师，并且每个教师在名录中只能出现一次。

1. 抽样框应该包括的内容

一般说来，抽样框应该包括以下信息：

(1) 识别信息

识别信息指的是惟一的、用于识别抽样单元的信息，如姓名、准确的地址、身份证号码等。比如说，北京广播学院教师名录中一定要包括每个教师的姓名。

(2) 联系信息

联系信息是指和抽样单元进行联系的方式，如电话号码、通信地址等；比如说，北京广播学院教师名录中一定要包括每个教师的联系方式，可能是电话，也可能是通信地址或者是电子信箱等等。

(3) 分类信息

分类信息是指与抽样中分类抽选有关的数据，例如，如果抽样时要求男性与女性分别抽选，则抽样框中一定要包含每个人的性别；对北京广播学院的教师住房情况进行调查，如果想来了解不同职称的教师的住房情况，则抽样框中要有每个教师的职称情况。

对于一个抽样框来讲，识别信息与联系信息是必需的；但分类数据有助于提高调查的效率。

2. 抽样框的类型

在实际工作中，有两种类型的抽样框：目录框和区域框，其中目录框又可以分为实际目录框和概念目录框。

(1) 实际目录框

实际目录框是包括总体中所有个体单位的实际名录，比如

说，由工商部门公布的所有注册公司的名单。实际目录框，可以通过各种不同的来源获取。各种各样的机构以及不同级别的政府部门，因各自的需要，都收集、保存有相关的目录，比如，一个居民小区的住户名录，一个行政区域内的公司名录，学校里的学生名录，工厂里的雇员名录等等。很多的实际目录框，都可以通过公开的出版物获得。

(2) 概念目录框

概念目录框就按照给定的要求所描述的个体单位的集合。比如说，在6月1日，所有进入北京动物园参观的人，就是一个概念目录框。要注意的是，概念目录框涉及的总体，通常只有当调查正在进行时才能建立起来。

(3) 区域框

区域框是一种特殊的抽样框，被抽样总体中的个体单位是由地理区域构成的。

区域框主要适用于以下的两种情况：

第一，当调查本身就是地理性质的，比如，通过计量每平方公里的动物数以推算野生动物的数目；

第二，有时很难获得适当的目录框，不得不使用区域框。在社会调查中，随着时间的推移，总体不断地变化，就一个小城市的常住人口来说，出生、死亡以及迁移都会对总体产生影响，使得总体的目录框很快就变得过时。比较而言，区域框随时间而变化的速度比较慢，在地图上地理单元的位置具有惟一性，访问人员很容易找到。

3．抽样框的缺陷

在实际工作中，很难找到和建立一个十分理想的抽样框。通过各种渠道所获得的抽样框，总是存在这样或那样的缺陷。抽样框的缺陷表现在以下几个方面：



（1）不完全覆盖

不完全覆盖是指总体中的某些个体单位排除在抽样框之外。比如说，在北京广播学院的教师名录中，可能没有包括近两个月内分入或调进来的新教师。前面我们所提到的《文学文摘》的1936年总统选举预测失败，失败的原因本质上就是抽样框不完全覆盖，理想的抽样框应该包括所有选民，但《文学文摘》的以电话号码以及汽车牌照登记簿上的名单作为了抽样框。很显然，这个抽样框没有包括所有的选民，并且所包括的选民与没有包括的选民在投票意向上存在很大差异。

（2）过涵盖

与不完全涵盖正好相反，过涵盖是指抽样框中包含了本来不属于总体的单位。

比如说，有几个教师已经调离了北京广播学院，而在北京广播学院的教师名录中，他们的名字依然存在。

（3）重复单位

重复单位是指某些个体单位在抽样框中出现不止一次。比如说，要了解北京广播学院的毕业生的工作情况，研究对象是毕业的本科生和研究生，那么通过学生处获得了历年的本科毕业生的目录，通过研究生部门获得了历年的研究生毕业的目录，把两个目录结合起来，作为抽样框，这样的话，有些人在北京广播学院上的本科，又上的研究生，那么，这些人在抽样框中就出现了两次。通常，只有在实际调查过程中才能发现重复的个体单位。

（4）错误分类

错误分类是指抽样框中的分类变量取值错误。比如，一个男性被错误地归类为女性，或者一个零售公司被归类为批发公司，这会导致抽样设计的失效和不完全涵盖，这是因为当我们

仅仅抽取零售公司时，我们会错过很多被错误划分为批发公司的零售公司。

抽样框会对调查的许多步骤产生直接影响，比如，抽样框会影响数据收集的方法，如果抽样框没有提供电话号码，那么就不能进行电话调查。在利用抽样框开始抽样工作之前，必须先对抽样框进行全面检查，以尽可能地减少抽样框的缺陷。

五、概率抽样与非概率抽样

根据样本抽取方法的不同，抽样可以分为概率抽样（probability sampling ）和非概率抽样（non-probability sampling）两类；也可以分为概率抽样、准概率抽样（Quasi-probability sampling）和非概率抽样三类。

1．非概率抽样

非概率抽样并没有严格的定义，主要特点是样本的抽选是根据主观判断有目的、有意识地或根据方便的原则来进行的，而不是根据随机原则来抽选的。非概率抽样效果的好坏在很大程度上依赖于抽样者的主观判断能力和经验，不能够计算抽样误差，不能从概率意义上控制误差并保证推断的准确性。非概率抽样包括整群抽样、方便抽样、判断抽样、配额抽样、滚雪球抽样以及自愿者抽样等具体形式。

2．概率抽样

概率抽样是严格按照给定的概率来抽取样本的。即总体中的每个个体都有一个非零的被抽中的概率。

在民意测验中，如果调查对象是北京市所有 18 岁以上的成年人，则概率抽样要保证每个合乎条件的人被抽中的机会相同。

概率抽样包括简单随机抽样、分层抽样、PPS 抽样、系统抽样和多阶段抽样等具体形式。



3．采用概率抽样还是非概率抽样

确定一项研究是采用概率抽样还是非概率抽样，要考虑以下4个方面的因素：

（1）研究目的

一些调查研究不需要推断总体，只是了解变量之间的关系，或是为问卷设计及测量尺度收集探索性的数据。这种情况下，常常采用非概率抽样的方法。

在选举预测中，关心的是全体选民的投票结果，在收视率研究中，关心的全体电视观众的收视行为，所以要采取概率抽样的方法。

（2）成本与价值

抽样应该用最小的投入产生最大的价值，如果对于概率抽样来讲，成本太高，则常常选择非概率抽样。

（3）时间要求

很多时候，一项研究是有时间限定的，而概率抽样相对说来，需要的时间会长一些．这时候，也常常选择非概率抽样。

（4）可接受的误差大小

在一些初步的或探索性的研究中，误差的控制不是严格要求的，则可以考虑采用非概率抽样的方法。

六、抽样误差与非抽样误差

在有关调查的问题中，人们经常问："被访者告诉的是真话吗？"这个问题的答案就是："如果他们知道答案的话，大部分时候被访者告诉的是真话。"[1] 这一点在有关总统选举的预

[1] Norman M. Bradburn & Seymour Sudman《Polls & Surveys-Understanding what they tell us》，Jossey－Bass Publishers，1988年，第8页。

测调查中得到了验证。当然，在某些敏感性问题上，被访者所说的不一定是真话，如有关吸毒、漏税、贪污、犯罪行为以及灰色收入等等。有时候，被访者由于记忆错误等原因，也会没有告诉真相。

误差（error）是指调查得到的结果与客观存在的真实结果的差异，只要是调查就一定有或大或小的误差存在，不可能完全避免误差。在抽样调查中，误差可分为非抽样误差（non-sampling error）和抽样误差（sampling error）两类。

1．抽样误差

抽样误差是由于抽样造成的误差，是由于用样本估计总体而产生的误差，样本只是总体的一部分，用局部数据作为整体的估计不可能完全正确。对任何一种抽样方案，可能的样本会有很多，而实际抽到的只是一个样本。因此，抽到哪一个样本完全是随机的、偶然的。如果抽到的是另一个样本，对总体的估计就会不同。这就是抽样误差产生的根本原因。只要是抽样调查，这个误差是不可避免的。

抽样误差的大小取决于很多因素，包括样本容量的大小、总体的大小、所感兴趣的总体特征的变异程度、抽样设计以及估计方法等，详细的推导和说明请见专门的抽样调查方面的教科书。

但与非抽样误差不同的是，对于概率抽样而言，抽样误差是能够计量的，而且可以得到控制的，只要愿意，可以将抽样误差控制在任意小的范围内。这也是概率抽样被采用的最重要的原因。

对于简单随机抽样而言，95%置信度下的二项变量的抽样误差如表3-3所示。二项变量是指问题的答案只有两个选项，如“是”、“否”；“同意”、“不同意”；“支持”、“反对”等。第



一列为样本量的大小，第一行和第二行表示简单随机样本所得到的二项变量的比例结果。

表 3-3　95%置信度下的简单随机抽样的抽样误差(%)

	1%	5%	10%	15%	20%	25%	30%	35%	40%	45%	
	99%	95%	90%	85%	80%	75%	70%	65%	60%	55%	50%
25	4.0	8.7	12.0	14.3	16.0	17.3	18.3	19.1	19.6	19.8	20.0
50	2.8	6.2	8.5	10.1	11.4	12.3	13.0	13.5	13.9	14.1	14.2
75	2.3	5.0	6.9	8.2	9.2	10.0	10.5	11.0	11.3	11.4	11.5
100	2.0	4.4	6.0	7.1	8.0	8.7	9.2	9.5	9.8	9.9	10.0
150	1.6	3.6	4.9	5.9	6.6	7.1	7.5	7.8	8.0	8.1	8.2
200	1.4	3.1	4.3	5.1	5.7	6.1	6.5	6.8	7.0	7.0	7.1
250	1.2	2.7	3.8	4.5	5.0	5.5	5.8	6.0	6.2	6.2	6.3
300	1.1	2.5	3.5	4.1	4.6	5.0	5.3	5.5	5.7	5.8	5.8
400	0.99	2.2	3.0	3.6	4.0	4.3	4.6	4.8	4.9	5.0	5.0
500	0.89	2.0	2.7	3.2	3.6	3.9	4.1	4.3	4.4	4.5	4.5
600	0.81	1.8	2.5	2.9	3.3	3.6	3.8	3.9	4.0	4.1	4.1
800	0.69	1.5	2.1	2.5	2.8	3.0	3.2	3.3	3.4	3.5	3.5
1000	0.63	1.4	1.9	2.3	2.6	2.8	2.9	3.1	3.1	3.2	3.2
2000	0.44	0.96	1.3	1.6	1.8	1.9	2.0	2.1	2.2	2.2	2.2
5000	0.28	0.62	0.85	1.0	1.1	1.2	1.3	1.4	1.4	1.4	1.4

由上表可知，在样本量为 1000 的电视收视率调查中，对于是否收看了某个节目，如果调查结果表明有 30%的人收看

此节目，则我们有95%的把握说，这个节目的收视率在27.1%（30%－2.9%）和32.9%（30%＋2.9%）之间。

2. 非抽样误差

非抽样误差是在抽样调查中由于人为的差错所造成的误差。非抽样误差产生的主要原因有以下几个方面：

（1）涵盖误差

涵盖误差也叫抽样框误差，是指在抽样框中出现了单位的遗漏、错误的包含、重叠以及错误的分类等缺陷。涵盖误差能影响调查中的每个估计，所以是一种极其重要的误差类型。

（2）测量误差

计量误差是由于对一个问题所做的记录与其真值不同而产生的。这可能是由受访者、调查员、调查问卷、收集数据类型或计量工具造成的。专业术语的使用、概念模糊、措辞不当等都可能影响被访对象和调查员对问题的理解；在直接测量中，所使用的测量工具也可能有误差存在。比如，在进行居民体重调查时，用于记录被抽中居民体重的台秤，可能因为刻度的校准有误而显示为不正确的值。

（3）无回答误差

无回答可分为两种类型：全部无回答和部分无回答。全部无回答，也叫单位无回答，是指被调查对象拒绝接受访问，没有提供任何信息。部分无回答，也叫项目无回答，是指被调查对象接受了访问，但是对问卷中的某些问题没有给予回答，是空缺的。

拒绝接受访问的人可能具有不同于一般受访者的特征，如果拒访率过高的话，调查结果的代表性将值得怀疑。与拒访率相对的是访问的成功率，有资料表明，在美国，由美国普查局（The Census Bureau）所做的人口普查（Current Population Survey）



的成功率为97%，在其他的一些以家庭为单位的调查中，成功率在90%以上；一般说来，被访者文化程度越高，访问成功率越高；其他一些机构所做的调查中，以家庭为访问对象的，成功率在80%左右，以个人为访问对象的，成功率在70%左右；固定样组的调查的成功率在50%以下。[①]

问卷问题过多、提到的概念很难理解或者访问员访问技巧太差，对调查的目的以及可能的用途解释不当，或不加解释都会使受访者失去信心，从而引起无回答误差。

(4) 数据处理误差

数据收集过程完成以后，就进入了数据处理过程，包括数据的编码、录入、整理等。在这些过程中，也可能产生错误。比如，在把问卷的内容录入电脑的过程中，某个问题的选项有5个，分别用1、2、3、4、5来表示，某份的问卷的答案是5，但录入员在录入过程中，把5录成了4，就产生了数据处理错误。

非抽样误差是无法测量的，只能通过一定的措施尽量减少它们的出现，比如，加强对调查人员的培训、加强管理、采用正确的调查方法、设计好问卷等等。

第三节　概率抽样的几种基本形式

在概率抽样中，总体中的每个个体单位都有一个非零的进入样本的概率。概率抽样最主要的优点就是用一个相对小的样本，可以对一个大的总体进行推断。

---

① Norman M. Bradburn & Seymour Sudman《Polls & Surveys-Understanding what they tell us》, Jossey-Bass Publishers, 1988年，第123页。

以下所介绍的是几种基本的概率抽样方法，实际调查中所采用的抽样方法常是这些基本概率抽样方法的组合。

一、简单随机抽样

简单随机抽样（simple random sampling）也称为单纯随机抽样，从总体N个单位中抽取n个单位作为样本，抽取方法是从总体中逐个不放回地抽取单位，每次都是在所有未进入样本的单位中等概率抽取的。准确地说，是样本中的每一个单位都有相同的入样概率。

简单随机抽样也可以一次同时从总体中抽得，只要保证全部可能的样本每个被抽中的概率都相等。

在日常生活中经常采用的抽签方法就是简单随机抽样。即把总体中的每一个单位都编上号码，并做成签，把所有的签放在一个容器里搅拌均匀，然后从中抽取一部分，这部分签所对应的个体单位就构成了样本。

在实际应用中常常利用随机数字表来实现简单随机抽样。随机数字表是由一些任意的数字毫无规律地排列而成的数字表。

随机数表的使用很简单，基本步骤如下：

第一，确定数字的位数，位数与总体中个体的数目所占位数相同，若总体为北京广播学院的所有学生，人数为6320人，则位数为4位；若总体为北京广播学院新闻系98级的所有学生，人数为40人，则位数为2位；

第二，在随机数字表中确定一个随机起点，可以是任意行、任意列；

第三，确定查找的顺序规则，可以是从上到下，也可以是从左到右；



第四，按照确定的原则进行随机抽样，符合条件的编号对应的个体就进入样本。

我们看一个具体的实例，假设我们想从98新闻班中抽取6个人去参加某项活动，总人数为40人，每个人对应着从01到40的学号，所以确定的位数为2位；在随机数字表中随机确定一个起点，比如说，是第4行第79列和第80列，查找的顺序是从上到下，则随机数字表上第一个数为53，超过最大学号，跳过；第二个数83，跳过；第三个数为48，跳过；第四个数为38，表示学号为38的同学被抽中；下一个数为99，跳过；下一个数为23，表示学号为23的同学被抽中。依次类推，直到抽出了6个同学。

表3-4　随机数字表（第一行为列号，第一列为行号）

	50~54	55~59	60~64	65~69	70~74	75~79	80~84	85~89	90~94	95~99
00	59391	58030	52098	87024	87024	82848	04190	96574	90464	29065
01	99567	76364	77204	27062	27062	96621	43918	01896	83991	51141
02	10363	97518	51400	98342	98342	61891	27101	37855	06235	33316
03	86859	19558	64432	99612	99612	59798	32803	67708	15297	28612
04	11258	24591	36863	31721	31721	94335	34936	02566	80972	08188
05	95068	88628	35911	33020	33020	80428	39936	31855	34334	64865
06	54463	47237	73800	36239	36239	71824	83671	39892	60518	37092
07	16874	62677	57412	31389	31389	62233	80827	73917	82802	84420
08	92494	63157	76593	03505	03505	72389	96363	52887	01087	66091
09	15669	56689	35682	53256	53256	81872	35213	09840	34471	74441
10	99116	75486	84989	23476	52967	67104	39495	39100	17217	74073
11	15696	10703	65178	90637	63110	17622	53988	71087	84148	11670
12	97720	15369	51269	69620	03388	13699	33423	67453	43269	56720

13	11666	13841	71681	98000	35979	39719	81899	07449	47985	46967
14	71628	73130	78783	75691	41632	09847	61547	18707	85489	69944
…	……	……	……	……	……	……	……	……	……	……
45	12900	71775	29845	60774	94924	21810	38636	33717	67598	82521
46	75086	23537	49939	33595	31484	97588	28617	17979	70749	35234
47	99495	51434	29181	09993	38190	42553	68922	52125	91077	40197
48	26075	31671	45386	36583	93459	48599	52022	41330	60651	91321
49	13636	93596	23377	51133	95126	61496	42474	45141	46660	42338

简单随机抽样的优点：

简单随机抽样理论完善，有标准的公式可以计算样本量的大小、总体估计和方差估计。所以是其他抽样方法的基础。当总体单元数不太大时，简单随机抽样实施起来并不困难。简单随机抽样中的估计方法，通常是采用样本均值来估计总体均值，用样本比例来估计总体比例，这就是简单估计。

简单随机抽样的缺点：

如果实际应用中，如果总体单元数 N 比较大，简单随机抽样实现起来就比较困难了。首先是不容易得到包含全部 N 个单元的抽样框；另外，利用简单随机抽样得到的样本单元可能比较分散，不利于调查访问的实施。还有一点就是简单随机抽样没有利用总体中已知的一些信息，不是最有效率的。因此，在实际中直接采用简单随机抽样的情况并不多。

二、分层抽样

把总体按一定的原则分成若干个子总体，每个子总体称为层，在每个层内分别抽取一定数量的单元构成样本，不同层的抽样相互独立，这样的抽样称为分层抽样（stratified sampling）。



1. 分层抽样的步骤

第一，确定分层指标。要根据研究目的的不同，选择合适的分层指标。分层指标可以是一个，也可以是多个。分层后，层内个体单位之间的差异要小，而使得层间的差异要大；

第二，按照分层指标，把总体分成 K 个互不重叠的部分，每一部分叫一个层；

第三，确定每层应抽取的样本量；

第四，在各层中分别抽取相应数量的个体单位，构成子样本，各个子样本的和为总样本。

假定某高校共有本科学生 1 万人，现想从中抽取 100 人来了解一下每月的生活费用问题，入样概率为 1%。采用分层抽样的方法，把年级作为分层指标，每个年级作为一层，共分为 4 层。一年级共有 3500 人，从中抽取 35 人进入样本；二年级共有 3000 人，从中抽取 30 人进入样本；三年级有 1800 人，抽取 18 人；四年级有 1700 人，共抽取 17 人。每层所抽取的样本相加，就构成了总的样本，共 100 人。模拟示意图如表 3-5：

表 3-5

总体	N 10,000			
层	N1 3500	N2 3000	N3 1800	N4 1700
子样本	n1 35	n2 30	n3 18	n4 17
样本	n 100			

2. 分层抽样的优点

(1) 抽样精度比较高：分层抽样充分利用了总体已有的信息，样本的代表性要强；

(2) 数据处理比较方便；

(3) 组织实施也比较方便，样本分布比较均匀。

3. 分层抽样的缺点

分层是根据研究目的，按照一定的指标来进行的，在抽样框中必须要具备用来分层的信息，才能采用分层抽样的方法。也就是说，分层抽样对抽样框的要求高。

三、PPS 抽样

简单随机抽样是等概率抽样，适用于总体中各个抽样单元的大小基本相同的情况。

如果总体中各个单元的大小差距比较大的时候，比如说，从全国所有大学中抽取 30 所大学进行调查，如果仍然采用等概率抽样的话，所取得的样本的代表性就会比较差。所以一般采用不等概率抽样方法。根据单元被抽中的概率以及是否有放回抽样，不等概率抽样可以有多种抽取的方法。其中最简单、最常用的方法就是有放回的按与抽样单元规模大小成比例的概率来抽取的方法，简称为 PPS（sampling with probability proportional to sizes）抽样。这里所说的规模大小是根据研究目的所确定的一个度量标准，可以是抽样单元的人数、单元的级别、重量、面积等等。

我们来看一个实施 PPS 抽样的实例：

中国互联网络信息中心在 2000 年底对有住宅电话的居民进行了上网情况的电话调查，每个省内要抽取 5 个地区，抽样指标为“住宅电话数目”，PPS 抽样，使每个地区被抽中的概

率，等于该地区"住宅电话的数目"与该省"住宅电话的数目"之比。下面我们以河北省为例，看一下具体的操作过程：

第一，对河北省"住宅电话数目"求和；

第二，确定每个地区"住宅电话数目"占全省"住宅电话数目"的比例及累计百分比，使总和为100%；

第三，用EXCEL软件产生0~1之间的均匀分布的随机数，根据随机数落在各地区对应累计百分比的范围，确定被抽中的地区；

第四，如果一个地区被抽到两次以上，则该地区样本量相应加倍。

生成的5个随机数分别是：0.512、0.275、0.427、0.622和0.913，则被抽中的5个地区分别是：邢台、秦皇岛、邯郸、保定和廊坊。见表3-6：

表3-6

地区	住宅电话数目	百分比	累计百分比	是否被抽中
石家庄市	1,261,180	0.140	0.140	
唐山市	1,079,711	0.120	0.260	
秦皇岛市	590,535	0.066	0.326	√
邯郸市	924,854	0.103	0.429	√
邢台市	755,743	0.084	0.513	√
保定市	1,099,658	0.122	0.635	√
张家口市	658,630	0.073	0.708	
承德市	573,773	0.064	0.772	
沧州市	800,608	0.089	0.861	
廊坊市	641,189	0.071	0.932	√
衡水市	609,199	0.068	1.000	

PPS 抽样优点：

PPS 抽样可以提高效率。

PPS 抽样缺点：

PPS 抽样对总体参数的估计公式比较复杂，另外，对抽样框的要求也比较高，要具备可以用来度量单位规模的信息。

四、系统抽样

系统抽样（systematic sampling）也称为机械抽样。如果总体中的单元都按一定顺序排列，在规定的范围内随机地抽取一个单元作为初始单元，然后按照一套事先确定的规则来确定其他样本单元，这种抽样方法称为系统抽样。严格意义上讲，系统抽样不是概率抽样，而是准概率抽样。

等距抽样是最简单、最常用的系统抽样，即取得一个初始单元后，按相等的间隔抽取样本单元。系统抽样最主要的优点是实施简单，因为只有一个初始单元需要随机抽取。

实施步骤如下：

第一，确定抽样间距：如在 40 人中抽取 6 人作为调查对象，则抽样间距 K 为：

K = N/n = 40/6 = 6.6 = 7

第二，确定随机起点：随机起点可以限制在第一个间距之内，也可以不做限制。

如，限制随机起点在第一个间距以内的话，即为 1～7，随机确定起点为 2，则 2、9（2+7）、16、23、30、37 对应的个体被抽中。



1	[2]	3	4	5	6	7	8
[9]	10	11	12	13	14	15	[16]
17	18	19	20	21	22	[23]	24
25	26	27	28	29	[30]	31	32
33	34	35	36	[37]	38	39	40

如，随机起点可以在1～40之间的话，假定随机确定起点为25，则25、32(25+7)、39、6、13、20对应的个体被抽中。

1	2	3	4	5	[6]	7	8
9	10	11	12	[13]	14	15	16
17	18	19	[20]	21	22	23	24
[25]	26	27	28	29	30	31	[32]
33	34	35	36	37	38	[39]	40

有时候，系统抽样甚至不需要一个完整的抽样框，只需要总体单元的一个确定的排列顺序就可以了。如果对总体单元的排列规则有所了解并加以正确利用的话，系统抽样能达到相当高的精度。

系统抽样的优点：

实施起来非常方便、易于操作。

系统抽样的缺点：

系统抽样最主要的缺点是估计量的精度估计比较困难。事实上许多行之有效系统抽样并不是严格的概率抽样。

周期性：要注意间距不要与事物本身的固有周期相重合。比如说，交通流量的调查，如果确定每7天调查一次，则起点是星期一的话，所有的调查时间都是星期一；如果起点是星期

日的话，则所有的调查时间都是星期日。对于一个城市来讲，星期一的交通流量是不能代表一周的情况的，同样，星期日的调查结果也不能代表一周的情况。

五、多阶段抽样

多阶段抽样（Multistage sampling）在大规模的社会调查中应用十分广泛。特别是当抽样单元是各级行政单位时，一般都采用多阶段抽样。

例如，在全国收视率调查中，先抽省，然后从抽中的省中抽取市/区，然后抽街道/县，然后是居委会/村，最后是户。见表3-7：

表 3-7

			中国		
			↓		
第一阶段	……	黑龙江省	北京	河北省	……
			↓		
第二阶段	……	西城区	朝阳区	海淀区	……
			↓		
第三阶段	……	朝外	管庄	呼家楼	……
			↓		
第四阶段	……	长营村委会	北京广播学院家委会	定福庄南里居委会	……
			↓		
第五阶段	……	李四家	张三家	王五家	……



多阶段抽样中由于每一级抽样中可能采取不同的方法，最后抽样误差的计算就十分复杂，所得到的样本不能看做是简单随机样本，不能直接应用现成的统计软件包（一般的统计软件，如 SPSS、SAS 要求总体中的个体被抽取的可能性相同，在这个前提下，对总体的参数进行估计和检验）。

在大规模的调查中，采用多阶段抽样，在各个阶段对抽样方法加以限定，以保证最终获得的样本可以近似看成是简单随机样本，也称为自加权样本。

多阶段抽样的优点：

在实际工作中经常使用，可以节省人力、物力。

多阶段抽样的缺点：

抽样误差相对比较大。

第四节　非概率抽样的几种基本形式

一、整群抽样

整群抽样（cluster sampling）就是先将总体中的各个单元归并为数量较少而规模较大的单元，称为群，抽样仅是抽群，对抽中的群调查其中所有的个体单元，对没有抽中的群则不需要进行任何调查。

比如说，某个学校共有 6500 学生，分布在 130 个班中，平均每班有 50 人左右，现想调查 250 人，则采用整群抽样的话，就是先从 130 个班中，随机抽取 5 个班，对抽中班级中的所有同学都进行调查。见表 3-8：

表 3-8

总体	N 6500				
群	R1 53	R2 50	R3 58	……	R130 45
抽样	R1 53	R4 48	R33 52	R98 50	R130 45
样本	n 250				

在整群抽样中并不需要提供包含每个次级单元的抽样框，只需要关于初级单位的抽样框就可以了。在上例中，我们不需要知道全校所有的学生名录，只要知道全校所有的班级的名录就可以进行抽样，对抽中的班级，调查所有的学生。

整群抽样的优点：

整群抽样最主要的优点就是实施方便，由于群通常是由那些地理位置邻近的或属于同一个系统的单元构成的，所以整群进行调查大大有利于实施，可以节省费用。

整群抽样的缺点：

整群抽样的精度较差，效率不高。因为同一群中的单元或多或少地有些相似，因此对抽中的群每个小单元都进行调查就不可避免地造成浪费。

二、方便抽样

方便抽样（convenience sampling）是选择最容易访问的人或最容易取得的资料作为研究样本。如，可以在调查公司周围来选择被访者，可以在人来人往的街头选择被访者。

按方便抽样所获得的样本，无法判断抽样误差，不能推断



总体。因此，一般适用于初探性或预测性的研究。如产品的广告测试等。

方便抽样的优点：

方便抽样可以节省费用、实施方便、速度快。用最少的费用，在最短的时间内完成访问。

方便抽样的缺点：

方便抽样的抽样误差大，调查结果的代表性差。或者说，用方便抽样的样本对总体进行推断的话，风险很大。

三、判断抽样

判断抽样（Purposive sampling）也叫非代表性的配额抽样（Unrepresentative quota sample），是根据对总体的了解，事先设定一些标准，然后根据这些标准来筛选样本，这样获得的样本，所代表的不是一般大众，而是特定的人群。

比如说，想要了解消费者对某品牌的啤酒的评价，可以选择经常饮用啤酒的消费者作为访问对象。

在美国的选举研究中，研究人员常采用判断抽样的方法，在所有的选区中选出少数具有代表性的选区进行研究。比如说，研究人员根据过去多次选举的结果，发现有3个区的投票结果，与全州的投票结果非常接近，那么，研究人员就选择这3个选区进行调查，并以此来预测全州的选举结果。

判断抽样的优点：

研究人员可以根据自己的技能、知识和经验去选择被访者，抽样过程简单。

判断抽样的缺点：

容易因研究人员的主观判断偏差而导致严重的误差。

四、配额抽样

配额抽样（quota sampling）是根据研究目的的不同，按照总体的某些特性，来抽取样本。比如说，根据总体中性别、年龄、文化程度等特征的分布，使样本中这些特性的分配同总体相一致。但单位具体是如何抽取的，不作限制。

配额抽样又可以分为独立控制配额抽样和交叉控制配额抽样。

1．独立控制配额抽样

假设某项研究要调查200人，根据以往的资料已知总体中男性与女性的比例是相同的，年龄分段的比例是“30岁以下”占20%，“31～40岁”占30%，“41～50岁”占35%，“50岁以上”占15%。

则在抽样时，男、女各调查100人。

性别	人数
男	100
女	100
合计	200

各个年龄段调查的人数分别是：

年龄	人数
30岁及以下	40
31～40岁	60
41～50岁	70
51岁及以上	30
合计	200



2．交叉控制配额抽样

交叉控制配额抽样是指同时满足几个条件的交互分配比例。比如，已知总体中男、女比例相同，并且知道男性中“30岁以下”占20%、“31～40岁”占34%、“41～50岁”占28%以及“51岁及以上”占18%，同样已知女性中各年龄段的人数比例，如果要调查200人的话，则性别和年龄的分配比例具体如表3-9所示：

表3-9

年龄	男	女	合计
30岁及以下	20	20	40
31～40岁	34	26	60
41～50岁	28	42	70
51岁及以上	18	12	30
合计	100	100	200

配额抽样的优点：

配额抽样简单易行，可以保证总体的各个类别都包括在所抽样总体中，使得所抽取的样本不太偏重某一阶层或地区。

配额抽样的缺点：

在配额抽样的每个类别中，被访者是由访问员主观选取的，会有偏差。在1948年美国总统选举预测调查中，采用的就是配额抽样的方法，结果预测错误。

五、滚雪球抽样

顾名思义，滚雪球抽样（snowball sampling）就像滚雪球一样，通过少量的个体单位来找到更多的个体单位，从而获取信

息，前提是抽样单元之间有一定的联系。

比如，我们要对安徽省在京的小保姆的所思所想进行调查，由于很难得到合适的抽样框，一开始可以通过身边的亲戚朋友先找到几个小保姆，然后再通过她们来介绍认识的其他保姆，依次类推，就像滚雪球一样，使得找到的保姆越来越多。

滚雪球抽样的优点：

滚雪球抽样可以有针对性地找到被调查者。

滚雪球抽样的缺点：

滚雪球抽样是通过一定的关系找到的被调查者，被访者之间同质性可能比较大，可能得到有偏差的样本。

六、自愿者抽样

自愿者抽样（volunteer sampling）是指被研究对象都是自愿参加的，一般通过征募方式进行征集自愿者。比如，在医学实验上，为验证某种新药的疗效，需要一些患者来参与实验，那么，出于道德的考虑，一定要征得患者的同意，也就是说，患者是自愿参加实验的。

在报道民意调查结果时，自愿者调查经常被认为是假冒调查（pseudo-polls），一般认为自愿者抽样调查结果是无价值的、无用的甚至会误导受众的。也有人称之为“脏水”SLOPS（Self-selected Listener-Oriented Public Opinion Surveys），SLOPS脏水是指这类调查缺少可信度。

最常见的SLOPS就是打进电话调查，多数情况下，一个问题被设定为用“是”或者“否”来回答，公布几个电话号码，用某个键代表“是”，某个键代表“否”，打入电话者，只要按照自动的语音提示，按相应的按键就可以表示自己的回答。

这类SLOPS调查结果非常具有震撼力，因为通常回答者



非常多，成千上万。它经常也是一个主要的税收来源，因为每个打入的电话是要付费的。

另一种常见的 SLOPS 形式就是寄回问卷调查，某些报纸或杂志随刊刊登一份调查问卷，读者可以通过邮寄或者是传真的方式寄回问卷。

由于回答者的数量非常巨大，常常给人一种可靠的假象，如果 SLOPS 的结果来自 200 万人，似乎很难忽略。

SLOPS 调查的最大问题是你不知道调查结果对总体的代表性如何。在概率抽样调查中，我们知道每个样本的入样概率，所以我们可以推算调查结果对总体的代表性。在 SLOPS 调查中，我们做不到这一点。

在 1987 年，ABC 新闻通过两种方式调查了这样的一个问题："联合国总部是否应设在美国?"（Should the United Nations headquarters stay in the United States?）一种方式是设立了免费电话，请受众自愿打进电话表达意见，同时也采用概率抽样的方式进行调查。结果如表 3-10：①

表 3-10

调查方式	自愿打入电话（Phone-in Poll）	科学抽样（Scientific Poll）
人数	100,000+	500+
是(Yes)	33%	66%
否(No)	67%	25%
不知道(Don't Know)	0%	9%

① Sheldon R. Gawiser & G. Evans Witt《A Journalist's Guide to Public Opinion Polls》Praeger Publishers，1994，第 101 页。

另一个例子是在1980年，美国总统候选人卡特（Carter）和里根（Reagan）电视辩论后，ABC新闻设立了打入电话的调查，如果支持卡特的话，拨一个电话号码，如果支持里根的话，拨另一个号码，每拨一个号码要花50美分，共有72.7万个电话打入，结果表明里根与卡特的支持率为2:1。而同时进行的一项概率抽样的调查结果表明，里根与卡特的支持率不相上下。

选举过后，来自卡特和里根竞选阵营的成员都承认，他们事先知道ABC新闻设立了打入电话之事，动员各自的自愿者去拨打相应的电话。这个自愿者调查结果只是表明：里根阵营的组织动员工作做得更好。

这个例子告诉我们，我们不知道SLOPS的结果说明了什么。参与人数众多并不能说明什么问题。1936年《文学文摘》的预测错误就是个典型的例子。

自愿者抽样的优点：

医学上的自愿者抽样，避免了道德上麻烦。报纸、杂志上的随刊问卷调查节省了经费。

自愿者抽样的缺点：

代表性差，对调查结果的使用一定要慎重。



第四章　内容分析法

在日常生活中，我们每天都在比较随意的程度上作着内容分析，我们阅读字里行间的信息，在已知信息的基础上寻找一般性、规律性。在社会科学研究中，可以使用正规的方法，更客观、更准确、更有说服力地进行内容分析。

传播学者 Schwengels 与 Lemert 在 1986 年用内容分析的方法把报纸对强奸案件的报道与同时期警方的记录作了比较，结果发现，新闻媒介对强奸案件的报道，不仅无法使民众提高警觉，反而可能对民众有误导作用。在警方的记录中，有 89%的强奸案件，罪犯是受害者熟悉认识的人，而在新闻报道中，却有 60%的案件，罪犯是受害者不认识的人；在警方记录中，近亲强奸的案件共有 21 件，上报的只有 3 件；强奸案件发生的地点，新闻报道与警方记录也有很大出人，警方记录显示，一半以上的强奸案，发生的地点是在罪犯或受害者家中，但在新闻报道中，却有 70%以上的案件发生在家庭以外的地点。

第一节　什么是内容分析

内容分析的应用起源于 18 世纪的瑞典，在 20 世纪 30 年

代，随着宣传分析和传播研究的发展，开始受到重视，近20年来，随着信息社会的到来，电脑的广泛应用，社会科学、人文科学日益重视定量方法，语言与文学研究也日渐科学化，内容分析研究更有了飞速的发展，目前，内容分析已经成为社会科学的重要研究方法之一，特别是在传播学领域，内容分析是常用的研究方法。

内容分析的研究对象十分广泛，可以用来分析任何形式的传播内容，包括书本、报纸、杂志、电视节目、电影、广告、录音带、诗词、歌曲、绘画、演讲、信件、法律条款、文件、记录等。

一、内容分析的特点

很多学者对内容分析下过定义，大同小异，即内容分析（content analysis）是一种客观的、系统的、量化的对传播所显示出来的内容进行研究的方法。内容分析的定义，强调了以下的几个特色：

1. 系统性

系统性（systematic）是指采用内容分析进行研究时，在选择研究样本方面，必须采用随机的抽样方法，而不是只选择那些支持研究者假设的资料；在建构类目与编码时，也必须遵循系统的程序，即类目的定义与编码要完全符合，始终如一。这样才能使研究结果具有推论能力。

2. 客观性

客观性（objective）是指在内容分析的研究过程中，每一步骤都必须制定一套明确的标准与规则，研究人员必须严格遵守，以排除人为的主观偏好与成见，并使不同的研究者能够按照同样的程序，验证研究结果。



3．定量性

定量性（quantitative）是指内容分析是一种量化的研究，在研究过程中，研究者必须对所有的变量下操作性定义，并决定测量变量所用的测量尺度，这样才能使内容分析的结果能够进行统计分析，从而更精确地解释研究发现。也就是说，按照既定的规则，对分析单位加以计量，用数字比较符号文字出现的次数，以达到准确的要求。

4．描述性

描述性（descriptive）是指内容分析方法的目的是描述传播的内容，即通过量化的指标对传播内容进行描述。在统计分析方面，所用到的方法以描述性统计分析为主，如频数、比例、众数等。

5．显明性

显明性（obvious）是指对传播内容所明显表现出来的内容进行分析，而不是针对隐含的意义进行分析。

内容分析方法特别适用于大众传播媒介的传播研究，并回答传播研究的经典问题："谁说了什么、对谁说、为什么说、如何说以及产生什么影响。"

二、内容分析的优点

1．节省人力、物力

一般说来，与调查研究法以及实验研究方法相比，内容分析方法不需要大量的工作人员，也不需要仪器设备，可以节省时间、费用和人力。

王石番教授（台湾国立政治大学新闻传播学院院长）在《传播内容分析法——理论与实证》一书的自序中写到："负笈美国明尼苏达大学新闻与传播学院时，修习传播研究方法，

师长常要求必须组成小组，从事实证研究，且限定只可从社会调查法、实验法和内容分析三者中选择一种作习题。初履斯土，人地生疏，加上语言、外表异于美国人，即刻进行社会调查研究，实有困难；如以实验法从事研究，招募实验对象，寻觅实验环境，也非易事。于是只得以成本最低的方式，借阅图书馆报刊杂志，制定编码表，与伙伴枯坐斗室，努力阅读，完成习题作业，虽然苦不堪言，却甘之如饴。在开始那一段在美求学的日子，内容分析普渡了我们这一组的熬夜生涯，终能过关。”

2. 研究范围广泛，可以超越时空限制

内容分析研究的是记录下来传播内容，只要记录的载体存在，就可以进行研究。可以研究现在的报纸，也可以研究 10 年前的报纸；既可以研究中国的报纸，也可以研究外国的报纸。研究的范围非常广泛，不受时空的限制。

北京广播学院柯惠新教授所主持的一项研究，内容是两岸三地的报纸对 1999 年台湾“9·21”大地震的报道差异。在内陆，选取的报纸是《人民日报》和《北京青年报》；选取的香港报纸是《明报》和《东方日报》；选取的台湾报纸是《中国时报》和《自由时报》。时期是 1999 年 9 月 16 日至 1999 年 10 月 15 日。

社会学者周敏在分析纽约地区唐人街华裔企业的变化和发展情况时，对电话号码簿进行了分析，“唐人街出版的电话号码簿列出纽约地区的大部分华裔企业，可查出唐人街内、外的华裔企业的类型和地址，还能从不同年份的电话号码簿看出这些信息随时间变化的情况。只要简单统计一下，就得到唐人街经济活动的分布情况；比较 1958 年和 1988 年的电话簿，就知



道华裔企业30年间的变化和发展趋势。"[1]

3. 非接触，可反复进行

非接触（unobtrusive），也叫做非介入，是指研究者不会直接接触到研究对象。采用内容分析，研究者探究人们制造出来的传播内容，而不是直接观察人们的行为，或者请人们回答问卷。在实验法和调查法中，研究的执行者和研究对象是互动的，是动态的，有时会使被研究对象心有顾虑而改变行为和回答内容。而内容分析方法是单方面的，对象是静态的，分析的是传播内容，是先有传播内容的，而后才进行分析，自然不会对传播者造成影响。

另外，在调查或实验方法中，如果事后发现调查和实验过程有缺陷存在，想进行补救是十分困难，甚至是不可行的。而在内容分析中，如果发现类目建构得不合理或者是编码有误，就可以重复进行。

三、内容分析的缺点

1. 只能研究记录下来的传播内容

内容分析所研究的对象必须是被记录下来的内容。有时候，研究者所关心的研究主题可能没有足够的资料，那么，就不适合采用内容分析方法。

2. 编码困难

在内容分析方法中，研究者要花费大量的时间来查看大量的资料，以确定类目，由于所研究的文献资料没有标准化的形式，因此编码和分析很困难。

---

① 周敏《唐人街—深具社会经济潜质的华人社区》，商务印书馆，1995年，第8页。

第二节　内容分析的功能

在精确新闻报道中，内容分析方法主要可以帮助完成以下功能：

一、揭发社会问题

揭发社会问题是记者最重要的职责之一，运用传统的新闻报道技巧，记者通常只能以具体的案例来探讨社会问题。由于具体的案例不具有推论能力，无法证明所报道的案例是普遍存在的社会问题，因此，价值有限，不易引起社会的关注。

内容分析方法能帮助记者以科学的方法整理文件记录，提出确切的证据，揭发普遍存在的或是潜伏已久的社会问题。

费城询问者报（Philadelphia Inquirer）的两位记者，用内容分析的方法分析了费城1971年发生的1374件刑事案件，结果发现：

• 当黑人和白人犯相同的罪时，黑人被判刑入狱的比例较高；

• 当黑人犯罪，白人是受害者时，黑人被判的刑比较重；

• 当白人犯罪，黑人是受害者时，白人被判的刑比较轻。

这项研究充分表明，种族不平等的现象仍然存在于费城的刑事司法制度中。探讨这种长期存在的社会问题，如果采用传统的报道方法，记者最多能举出少数案例，缺乏说服力；采用内容分析方法，就能提出确切的证据，使所报道的问题，受到社会各界的重视。

事实上，很多社会问题都可以用内容分析方法来加以探讨。



二、探讨社会趋势

社会发展趋势是非常有吸引力的题材，著名美国社会学家、未来学家奈斯比特（Naisbitt）1982年的畅销书《大趋势》（Megatrends）和1990年的《2000年大趋势》（Megatrends 2000），就是利用内容分析法，有规律地考察分析美国各地的报纸，每月多达数千份，从而探讨现代美国生活的主要趋势。

《大趋势》和《2000年大趋势》这两本畅销书给新闻界的启示是，社会趋势是社会大众关心的问题，而记者可以用内容分析来探讨社会趋势。负责社会新闻的记者，可以分析历年的犯罪记录，探讨近年来社会主要犯罪类型的转变；负责文教影视的记者，可以分析电影、电视节目、小说、报纸以及杂志的内容，从而探讨社会价值观念的变化等等。

三、评估媒介表现

采用内容分析的方法，即通过比较新闻界的报道和实际情况的差异，深入评价新闻界的整体表现，1986年，传播学者Schwengels与Lemert用内容分析的方法把报纸对强奸案件的报道与同时期警方的记录作了比较，结果发现，新闻媒介对强奸案件的报道，不仅无法使民众提高警觉，反而可能对民众有误导作用。

另外，在传播学研究领域，按照传统的传播模式（“谁”传递“什么”内容，给“谁”，用“什么”形式，达到“什么”效果）来分类，内容分析方法可用于检视传播内容本质、探究内容表达的形式、分析传播来源的特质、了解受众的特性以及验证传播内容的效果。台湾学者王石番在《传播内容分析法——理论与实证》（幼狮文化事业公司印行，1999年）一书

中，对内容分析方法的功能作了如下的归纳：

1. 检视传播内容本质	○描述传播内容的趋势 ○追溯学术研究发展的轨迹 ○揭示各国传播内容的不同 ○比较传媒处理论题的层次与内涵 ○依据既定标准，审核传播内容 ○建构传播媒介应履行的标准
2. 探究内容表达的形式	○揭示宣传技巧 ○衡量传播内容的可读性 ○发现题材特征
3. 分析传播来源的特质	○检视传播者的特征 ○测定个人和团体的心理状态
4. 了解受众的特性	○鉴定团体与人物的形象 ○反映群体的态度、兴趣、价值和文化类型 ○提供犯罪证据 ○获取政治和军事情报
5. 验证传播内容的效果	○揭示注意焦点 ○描述传播对态度和行为的影响 ○比较媒介内容与真实世界 ○建立媒介效果研究的起点



第三节　内容分析的研究程序

一般说来，采用内容分析法进行研究，具体的研究过程，可细分为以下步骤：

确定研究问题或假设→界定研究总体→抽取样本→界定分析单位→建构类目→编码→统计分析→解释结果

北京广播学院曹培鑫的硕士论文采用内容分析的方法分析了我国发行量最大的文摘杂志《读者》中“人物”栏目的20年变化的情况，目的是从中了解20年来社会价值观的发展变化（以下简称《〈读者〉价值观》）。下面，我们以此为例来对内容分析的每个步骤加以说明。

一、确定研究问题或假设

在具体的研究开始以前，首先要确定研究的问题是什么？研究的假设又是什么？在《费城询问者报》记者用内容分析的方法分析了费城1971年发生的1374件刑事案件的例子中，记者们的研究问题是刑事司法制度中是否存在着种族不平等的现象。

在《〈读者〉价值观》的研究中，研究的问题是“‘人物’栏目20年人物的变化如何？”

二、界定研究总体

界定总体指明确规划内容主体的界限，即赋予总体以适当的操作定义。如果你想分析流行歌曲的内容，必须界定流行歌曲的范畴，也就是必须指明什么样的歌曲才是流行歌曲。可以按照某个社团的定义，官方的定义等等，也可以自己参照有关

资料，自行定义。

在《〈读者〉价值观》的研究中，总体的界定比较清楚，20年来的《读者》杂志“人物”栏目的文章就是研究总体。

三、抽取样本

研究总体界定后，根据人力、物力的限制，要确定是采用全面了解还是抽样调查的方式。一般说来，由于人力、物力所限，多抽取样本作为研究资料。内容分析要特别考虑资料的分布情况，应该先把研究资料过目，熟悉资料结构，掌握总体特征。再决定抽样方法，使所抽样本具有代表性。

在《〈读者〉价值观》的研究中，对《读者》“人物”栏目从1981年第一期到2000年第二十四期的全部20年的文章共622篇进行了分析，即采用的是普查的方法。

四、界定分析单位

分析单位是内容分析过程中，必须实际计算的最小元素。分析单位随类目的不同有所不同。印刷内容常常以字、句、语干、段、篇等作为分析单位，而影视内容的分析常以人物、剧情等作为分析单位。

在《〈读者〉价值观》的研究中，分析单位是《读者》杂志中“人物”栏目的文章，即以篇作为研究单位。

五、建构类目

类目就是所研究内容的分类。建构类目就是把研究问题概念化和操作化。类似于调查研究方法中的问卷设计，所有的类目构成一份编码表，不同的是，调查问卷的研究对象是人，询问的是被调查对象的想法，而内容分析中，研究对象不是人，



而是各种传播内容，编码表的填写是由编码员根据要求来做出的判断。

在建构类目时，注意要符合研究类目，反映研究要求，同时各类目的选项应该既要穷尽又要互斥，另外类目的数量要适当，类目太少或许不能涵盖研究内容，无法达到详尽的要求，但类目太多，过于琐碎，又会影响资料的完整性。

在《〈读者〉价值观》的研究中，是从4个大的方面来建构类目的。即文章的基本资料、所记述的人物、主题、人物的追求与特征。详细的编码表和编码说明请见本章最后的应用实例。

在传播学研究中，类目的种类可分为“说什么”（what is said）和“如何说”（how is said）两大类，每大类又可以细分为一些小类。

说什么		如何说
◇主题	◇来源	◇传播形式
◇方法	◇目标	◇叙述形式
◇特性	◇标准	◇强度
◇主角	◇方向	◇策略
◇权威	◇价值	

简略说明如下：

1.“说什么”类目

(1) 主题类目

主题（subjects）类目指建构类目时，依据传播内容的主题加以分类。

例如台湾的学者在分析报纸的社论时，通常把社论的主题分为以下类别：

1. 元首； 2. 政府； 3. 议会； 4. 司法； 5. 军事； 6. 大陆问题； 7. 国际事务； 8. 执政党； 9. 其他政党； 10. 其他政治问题； 11. 财经问题； 12. 社会问题； 13. 文化活动； 14. 宣扬国家意识； 15. 其他。

（2）方法类目

方法（methods）指达到目的的手段。

例如，分析小说的故事情节，把小说中人物成功的方法，分成：

1. 机遇；2. 个人努力；3. 家庭背景；4. 投机取巧；5. 其他。

（3）特性类目

特性（traits）指传播内容来描述人物的特征，包括年龄、性别、种族（民族）、职业、婚姻状况、社会阶层等。

（4）主角类目

主角（actor）指传播内容中，各种场合的代表性人物或领导人物。

（5）权威类目

权威（authority）是指以人或团体的名义来发表声明或谈话。

例如，分析报纸所刊登的外电来源，把外电来源分成：1. 美联社；2. 合众社；3. 路透社；4. 法新社；5. 共同社；6. 其他。

（6）来源类目

来源（origin）指传播活动的发生地点。

如分析新闻时，依据新闻事件所发生的地点，把新闻分



成：1．国际新闻；2．国内新闻等。

(7) 目标类目

目标（target）是指传播的对象，传播常常以某些特定的人或团体为诉求对象。

(8) 标准类目

标准（standard）是以传播内容特性为标准加以分类的。

在分析电视广告中妇女的形象时，把电视广告中出现的妇女，依照性格分为乐观、悲观；独立、不独立；主动、被动；积极、消极等等。

(9) 方向类目

方向（direction）指传播内容所显示的态度与立场，常见的方向类目有赞成或反对、有利与不利、同意与不同意等等。

(10) 价值类目

价值（value）指传播内容显示出的哪些价值、目标或希望。

2．"如何说"类目

(1) 传播形式或类型

传播形式或类型（form or type of communication）很容易区分。在分析电视节目的形态时，把电视节目分为：1．新闻；2．教育文化；3．娱乐；4．公共服务；5．其他等。

(2) 叙述形式

叙述形式（form of statement）是指传播内容的文法或造句形式。

分析新闻报道时，把新闻报道的结构分为：1．报道；2．推论；3．意见等。

(3) 强度类目

强度（intensity）是指传播内容表达态度、感觉、立场的

强烈程度。

强度类目与方向类目相似，方向类目只是显示方向，但未指出态度、感觉或立场的强度；强度类目则不仅显示方向，也呈现强度。

(4) 策略类目

策略（device）是指传播内容中所采用的修辞或宣传方法。

分析电视广告的策略，把广告所采用的诉求方式分为：1. 感性诉求；2. 理性诉求；3. 恐惧诉求；4. 其他。

六、编码

按照既定的标准，将分析单位划归到具体类目的过程就是编码（coding），编码是内容分析中最费时费力、枯燥无味但是十分重要的工作。

为使同一编码员不同时间的编码过程一致，以及不同编码员的编码过程一致，必须设计标准的编码表（standardized coding sheet）。

编码员按照类目和分析单位的定义，阅读相应内容，判断分析单位应归入的类目，针对每个分析单位，填写编码表。

编码员的数目不宜过多，3 到 6 位比较合适。编码员必须先接受编码训练，以熟悉类目及其定义、规则后，才开始编码。一般说来，编码员的文化程度要求比较高，多具有大学以上文化程度。在学术研究中，多由在校大学生和研究生来担任。

在《〈读者〉价值观》的研究中，编码员是由曹培鑫的师弟、师妹来担任的，共有 6 位，其中有 4 位学习过内容分析方面的课程，曹培鑫对编码员进行了专门的培训。

如果不同的编码者按照既定的规则与定义来操作，但是编



码结果差距很大，则说明规则不够清楚，定义得模棱两可，即信度比较低。出现这种情况，必须停止编码，重新规划类目与分析单位的定义。一般说来，在正式编码前，对编码表进行预测和修改是不可避免的。

在《〈读者〉价值观》的研究中，正式编码前研究者与全体编码员进行过两次比较细致的讨论，编码前的试测信度达到0.96以后才开始正式编码的。

七、分析资料

内容分析得到的是定量化的数据，要用相应的统计分析方法来分析处理。最经常使用的描述性统计分析，如频数、比例等。有时，也可以用到一些多元统计分析的方法。

在《〈读者〉价值观》的研究中，用到的统计分析方法包括频数、比例（百分比）以及分类的频数和比例。

八、解释结果

根据统计分析结果，结合研究目的和假设，就可以进行解释了。

我们来看一个具体的例子：

社会学学者风笑天于1992年曾对1950年到1990年40年间中国妇女的形象及其变迁进行了研究，运用了内容分析的方法，选择了《中国妇女》杂志的“人物通讯”栏目的主人翁作为分析对象。根据研究目的，对每位主人翁都按照7个指标进行分类和编码。[1]

---

① 风笑天《社会学研究方法》，中国人民大学出版社，2001年3月，第224页。

<table>
<tr><td>(1)年龄</td><td>—30岁以下(青年)
—30岁～49岁(中年)
—50岁以上(老年)</td><td></td><td>(2)政治面貌</td><td>—党团员
—一般群众
—未提及</td></tr>
<tr><td>(3)文化程度</td><td>—小学及以下
—初中
—高中或中专
—大专以上
—不详</td><td></td><td>(4)行业</td><td>—工业
—农业
—商业服务业
—文教科卫
—机关部队
—其他</td></tr>
<tr><td rowspan="2">(5)职业</td><td rowspan="2">—工人
—农民
—商业服务业人员
—教师及医务人员
—文化艺术体育人员
—家属及离退休人员
—科技人员
—企业家及专业户
—军人及公安人员
—干部</td><td rowspan="2"></td><td>(6)劳模状况</td><td>—劳模代表
—非劳模代表</td></tr>
<tr><td>(7)事迹</td><td>—个人成就
—工作业绩
—服务他人
—人生经历
—贤妻良母
—用于斗争
—女强人</td></tr>
</table>

研究结果表明，五六十年代的女性形象多以年轻的、低文化程度的普通工人、农民为代表，以党团员、劳动模范和努力生产的业绩为特征；70年代末则以年老的、较高文化程度的文教科卫人员为特征，具有努力生产、个人成就、勇于斗争等多种形象特点；而80年代则以高文化的中年文教科卫人员为代表，以个人专业成就和女强人式的工作业绩为主要特征。

40年来总的趋势是：从五六十年代单一的劳模性模式，到70年代末的劳模型、个人成就型、勇于斗争型并存模式，



又到80年代以个人成就型和女强人型为主的模式。女性形象在不同时期所表现出的这些特点，从不同的侧面折射出我国社会各个历史时期的风貌，反映出整个社会历史的巨大变迁。研究结果还表明，女性角色形象是时代的产物，是漫长历史过程中的一种客观反映，是新中国成立40多年来我国社会中各种新事物、新观念、新思想与传统文化、传统观念交互作用，共同影响的结果。社会历史的发展不断改变着女性的角色形象，不断赋予女性形象以新的内涵，同时女性形象也不断地体现和反映我国社会各个时期的风貌。

第四节　信度与效度

一、信度

信度是指测量结果的可靠性和一致性。研究结果要有较高的信度，测量的方法与程序必须稳定可靠。也就是说，不同的研究人员采用同样的测量方法与程序，进行同样的研究，所获得的结果应该是相同的，具有可重复性；另外，在不同的时间，采用同样的测量方法与程序，进行同样的研究，所获得结果也应该是相同的。

就内容分析而言，研究资料的繁简难易、采用的类目、编码规则的明确程度、编码员的因素，包括技巧、经验以及对研究目的与设计的了解程度，都会影响信度。由于研究资料是客观存在的，品质的好坏是研究人员所不能控制的，所以，提高信度的方法只有改进编码员和类目的质量。

内容分析的信度分析具有3种类型：

1. 稳定性

稳定性（stability）指编码员的内在信度（intra-observer reliability）与一致性（consistencies），即同一编码员在不同的时间，依照相同的类目及编码规则，对相同的传播内容，进行两次编码，编码结果应该大致相同。

如果两次编码的结果相差很多，表示编码员本身不够稳定，或者是编码员对类目的界定以及编码规则不够了解。

一个编码员的稳定性，可以通过测试——再测试（test-retest）的方法测量，即让同一个编码员在不同的时间对相同的内容进行编码，然后比较两次的编码结果的异同。

2. 可复制性

可复制性（reproductivity）指编码员之间的信度（inter-coder reliability）与共识（consensus）。当不同的编码员在不同的环境，按照相同的类目与编码规则，针对相同的传播内容进行编码工作，编码的结果应该大致相同。如果编码的结果相差很多，则可能有两个原因：

第一，可能是类目的界定不够清楚，含义模糊，使得不同的编码员对类目的界定以及编码规则的理解不同。这时，研究者要对类目和编码规则加以修订。

第二，也可能显示编码员个人编码的过程不够稳定。

3. 正确性

正确性（accuracy）指内容分析的结果，符合某种既定标准的程度。

研究者先将部分传播内容，依照既定的程序，进行编码，在确定编码的结果正确无误时，便可把该编码结果作为评估的标准，来和编码员按照相同的程序，对同样的内容所完成的编码结果作一对照，从而测量编码员编码结果的正确性。



二、效度

效度是指研究结果的正确性，是指测验工具是否真正测量出所研究的现象，效度越高，表示分析结果越能显示所研究对象的真正特征。

内容分析的效度，一般分为以下 3 种：

1. 资料取向效度

资料取向效度（data oriented validity）是指分析方法能代表资料内信息的程度，又可分为语义效度（semantic validity）与抽样效度（sampling validity）。

语义效度主要是评估分析方法对资料内象征意义的敏感程度。当类目系统和测量方法的语义，符合传播来源、受众或其他相关资料的语义，就说明具备高度的语义效度。

抽样效度是指抽出的样本能正确代表总体的程度，样本的结构与总体的结构越接近，抽样效果越好。

2. 产品取向效度

产品取向效度（product oriented validity）是指分析方法在不同环境下的适用程度，又可分成相关效度（correlation validity）与预测效度（predictive validity）。

（1）相关效度目的在于评价用两种分析方法所得到的结果的相关程度；

（2）预测效度是指用一种方法所得到的结果，与直接观察的结果的吻合程度。

3. 过程取向效度

过程取向效度（process oriented validity）是指分析程序模式能代表资料脉络内关系的程度，这种效度又称为构念效度（construct validity）。

鉴定效度最传统的、最常用的、最有效的方法就是专家评审法，即由研究者邀请几位对所研究问题有独到见解的专家，请专家们对研究的设计、类目的界定以及抽样方法等研究过程做出评价。一般说来，在正式编码前，就应该先向专家进行咨询，听取专家的意见，改进研究设计的品质。

应 用 实 例

应用实例的编码表和编码说明选自北京广播学院传播研究方法方向2001年硕士毕业生曹培鑫同学的硕士毕业论文《〈读者〉的价值观评述——以对“人物”栏目的分析为例》。

《读者》编码表

编码前请再次详细阅读编码须知

总编号：____________

一、文章基本资料：

1．________年（1981～2000）________期（1～12，2000年1～24）________页（材料上的页码）

2．文章标题：____________

3．文章长度：约________字（每页2000字左右）

4．稿件来源：____________

5．原稿发表时间：________年

6．作者国籍编码：（请参看国家、地区编码表）________

7．体例：________

①自传　　②别人描写

8．文章主题：________



①情感 ②成就 ③性格 ④命运 ⑤道德 ⑥事件 ⑦轶事 ⑧观念 ⑨其他________

二、所记述的人物：

9．文章中主要记述了________个人物（此项为1则不用回答10、11题）

	身份	性别	国籍	评价
第一个				
第二个				
第三个				
第四个				
第五个				

身份、职业编码：

①政治家、官员、革命家	②军事家、军人	③文学家、学者
④画家、雕刻家、书法家	⑤音乐家(作曲)	⑥科学家
⑦演艺人员	⑧工人、农民	⑨经济、商业、金融、经管
⑩医生、医务人员	⑪教师、教练	⑫行政管理人员
⑬领袖人物	⑭社会工作者	⑮运动、武术
⑯单纯名人或与名人有关联	⑰怪才、奇人	⑱皇室成员、历史上大臣、大将
⑲侦探、警察	⑳航海家、探险家	㉑思想家、哲学家

㉒法官、律师	㉓记者、新闻从业人员	㉔宗教、神职人员
㉕学生、孩子	㉖建筑从业人员	㉗其他

性别：①男 ②女

国家、地区编码：

①中国	②美国	③法国	④苏联、俄国
⑤英国	⑥德国	⑦中国台湾	⑧意大利
⑨印度	⑩日本	⑪(古)希腊	⑫匈牙利
⑬西班牙	⑭比利时	⑮波兰	⑯挪威
⑰香港、澳门	⑱加拿大	⑲巴基斯坦	⑳新加坡
㉑荷兰	㉒瑞士	㉓瑞典	㉔奥地利
㉕爱尔兰	㉖以色列	㉗约旦	㉘泰国
㉙巴西	㉚南非	㉛埃及	㉜马其顿
⑬缅甸	㉞斯里兰卡	㉟菲律宾	㉟罗马尼亚
㊲尼日利亚	㉜丹麦		

评价：

①完全正面、崇拜	②正面评价为主
③不明显、难判断	
④负面评价为主	⑤完全负面、批判



10．人物之间的关系：________

①亲子	②家族	③朋友
④夫妻、爱人	⑤对立	⑥名人之间
⑦名人与普通人	⑧师生	⑨无明显关系
⑩其他（请注明）：		

11．人物之间的关系的性质：________

①同质关系　②异质关系　③混合关系

二、主题分述：

12．文章涉及的主要时间（只选一项）：________

①文革之后	②文革期间	③建国到文革之前
④1912 年～1949 年	⑤1840 年～1912 年	⑥1840 年之前
⑦跨时间段		

13．（12 题选 2 答此题）对于文革期间这一段时间的表述：________

①中性描述	②控诉	③反思原因
④忏悔	⑤未提到	⑥其他（请注明）：

14．（当第 8 题选 1 时答此题）情感的类型：________

①爱情　②友情　③亲情　④宗教感情　⑤其他________（请注明）

15.（当第8题选2时答此题）成就类型、领域：________

①经济型	②政治型	③学术型
④艺术、竞技	⑤道德型	⑥其他（请注明）：

16.（当第8题选3时答此题）人物的性格可以算做：______

①理想主义的　②现实主义的　③其他：________（请注明）

17.（当第8题选4时答此题）什么是决定人物命运的最主要因素：________

①政治因素	②经济因素	③人物性格
④人物观念	⑤时机	⑥其他（请注明）：

18.（当第8题选4时答此题）人物命运的主要方向：________

①积极正向的	②中性的无方向性的	③消极负向的

19.（当第8题选5时答此题）人物的道德可以归入：________

①传统的	②现代的	③其他（请注明）：



20．（当第 8 题选 6 时答此题）文中的事件是：________

①经济事件	②政治事件	③其他（请注明）：

21．（当第 8 题选 8 时答此题）观念种类：________

①经济	②政治	③学术
④生活	⑤艺术、竞技	⑥道德
⑦其他（请注明）：		

四、人物的追求与特征：

22．文中体现了主要人物对以下哪些方面的重视与追求？（最多可以选 6 项）

①舒适、自在的生活（赋予的生活）	②令人兴奋的生活（富于刺激、活跃）
③成就感（做出持久的贡献）	④和平的世界（没有战争和冲突）
⑤美丽的世界（自然和艺术的美）	⑥平等（四海之内皆兄弟，所有的人都赋予相同的机会）
⑦家庭安全（照顾自己所爱的人）	⑧自由（独立、有选择的自由）
⑨幸福（满足）	⑩内心的和谐（没有内心冲突）
⑪成人的爱（性和精神上的亲密）	⑫国家安全（不受攻击）
⑬快感（一个快乐、闲暇的生活）	⑭得救（被拯救而得永生）
⑮自我尊重（自尊）	⑯社会认可（尊敬、钦佩）



⑰真正的友谊(亲近的同伴关系)	⑱智慧(对生命有成熟的理解)

23. 文中的主要人物有以下哪些特征？(最多可以选6项)

①有抱负的（努力工作、有志向）	②胸开阔的（胸怀开朗）
③有能力的（有才能、有效力）	④欢愉的（轻松、欢快）
⑤干净的（整齐、井井有条）	⑥有勇气的（信念坚定、不怕困难、挫折）
⑦宽容的（愿意原谅他人）	⑧愿助人的（为他人的福祉而努力）
⑨诚实的（真诚、真实）	⑩富于想像的（大胆、有创造力）
⑪独立的（依靠自己、自给自足）	⑫智识的（聪明、内省的）
⑬有逻辑的（稳定、理性）	⑭有爱心的（热情、温柔、善良）
⑮服从的（顺从、尊敬）	⑯礼貌的（彬彬有礼）
⑰负责的（可依赖的、能依靠的）	⑯有自制能力的（拘谨、有自我约束力的）

㉔编码员姓名：____________

编码须知

研究简介：

此次研究的目的，是运用内容分析的方法，研究《读者》在其创刊发行的20年中，所传达的价值观的变化。内容分析



的文本，选取了杂志自创刊以来所有月份的期刊，从每个月份的期刊中选取了“人物”这一专栏进行分析。

一般注意事项：（请常常参考本说明）

1．总编号不用编码员编写。

2．每一个编码题选项中都有“不确定”，“不明确”，“未及”，如果归人此项，则数据的实用价值就不大了，所以请尽量将数据归人确切的选项，当然这也不是绝对的，要真实地记录。

3．编码中遇到的不能解决的问题，请找研究者协商。

4．很难定夺的问题、设计中未想到的问题，请在编码表的备注或者编码表背面做简要的说明，说明的时候请一定要注明无法归类的原因。

5．凡是选择“其他”选项的，请注明具体内容。

特别注意事项：

1．本编码表包括3个部分：

a）编码说明。

b）编码表详表。

c）录入表格。

其中编码详表是在编码之前对各个编码项的详细解释，在编码过程中的参照。熟悉了编码详表的结构之后，就可以进行编码了，同时将编好的数字录入到录人表格之中。

2．除了编码表中指定的不要在编码表上留空隙，不能填充的请注明原因，或“不详”。

3．字数：以500字为最小单位，每页按2000字计算，分成1/4页，1/2页，1页等。

4．对编码详表中各个题的解释：

A：

关于"主题"，分成了9种选项：情感、成就、性格、命运、道德、事件、轶事、观念、其他。主题可以理解成文章的中心，即文章在表达什么？有时文章可能同时表达了几个主题，这种情况下就要仔细思考、判断，将其归入其中最适当的一类。

情感	描写一人或者多人之间的感情为主，可以是父母之爱、情人夫妻之爱、同志感情等等以描写情感经历为中心。重点是通过事件描写感情。
成就	以描写各行业、领域的人物的成功经历为主，有时会描写从平凡走向成功的过程，或者描写成功过程，或叙述事业的进展，或者探讨成功的原因都归此类。
性格	或通过言论、或通过一系列事件而中心都是围绕着性格展开，各事件无不说明性格的方方面面。如果有一系列的事件，但并不围绕着一个固定的性格中心，则不能归入此类，往往要归入"轶事"类别中。
命运	此类文章常常会描写主人公的一生，多变起伏的命运，这样的文章不太关注主人公内在的部分，而是更加强调外在环境的变化及其影响，这种文章往往会描述一次以上的影响人物人生道路的事件。
道德	人物的一切行为的动因都是出于对高的道德标准和理想的追求，尤以利他主义为特征，常常得到赞美和起到精神鼓舞的作用。
事件	这类文章常常描述在历史事件中的人物，并不完全与人物为中心，文章常常具体详细地描写一两个事件，而事件本身的起因很可能是非人为的，所以事件本身成了重点。

轶事	常常是名人的生活中的小事，这些事如果不是因为事件的主人公是名人可能并没有什么记载价值，时间也并不反映人物很深层次的特征、常常是日常小事、小习惯等等。
观念	文章很少描述事件，或者简单的概括事件，而把重点放在了阐述表达人物的观念上面，这些观念可以包括各个方面的主题，往往是人物对于特定主题的看法，追求的理想等等。

B:

关于文章中出现的人物编码的方法：

1. “9题：人物个数”并不是所有出现在文章中的人都算在内，而是那些单独的构成了文章记述的部分人物，比如，“林肯和他的母亲”，“将军与诗人”的题目中的人物都是。还有“二战名将录”，“开国将军们”等描写完全相似的人物的文章都算描写了一个以上的人物。只有以上两种情况算做人物两个以上。

2. “10、11题：人物间关系（的性质）”只适用于两个以上的人物的情况，其中同质关系是指个人物的身份、职业、生活年代几乎完全相同（比如，“二战名将录”，“开国将军们”等），异质关系则是人物的身份很不相同，将几个人物放在一处不是因为相似，而是因为人物之间的关系（比如，“林肯和他的母亲”，“将军与诗人”等）。

C:

12、13题：文章涉及的主要时间，是指文章描述的主题部分的事件所涉及的时间段，如果很难归类则选7——跨时间段。

D:

14～21各题是分别对应情感、成就、命运、事件、观念的题，比如，在9题：文章“主题”时选了“情感”，就对第14题“情感类型”编码，其他的15～21则不用编码，另外几题以此类推。

其中：16题：现实主义的性格与理想主义的性格的分别在于，行动的标准是常常以现实状况为先决条件，还是往往不太考虑现实条件，以理想的状态为奋斗的目标。宗教的、乌托邦的，以及对各种主义的追求；反传统的，反权威的，往往是理想主义的性格特征；现世的、顺从的、保守传统的往往是现实主义的性格特征。

19题：传统的道德之中国以儒家为中心的道德体系所推崇的观念，包括：经济上的济贫救荒、政治上的德治教条、社会上的尊老敬贤、文学上的文以载道、宗教上的孝思观念等。现代的道德观念则是以西方道德框架为主的，自由、平等、博爱，尊重个体、提倡竞争等等观念。

22和23题：为人物的追求与特征，回答此题时不需要经过严密的、逻辑的推理，而应该将判断的标准停留在文章比较明显表达的地方或读者的直接感受。

E：

关于职业的编码：

1．各职业均为人物最主要的身份，比如，一个人既是军事家又是政治家时，要看他的军事才能是否比他的政治表现（比如，参加革命、担任政府官员等）更显著。

2．一个人的身份要与文章中的行为最接近的为准，比如，一个官员在从事文学创作中有很高的造诣，文章又正好在描述此事，则此时可能应该将其归入文学家类。

3．科学家：物理、化学、数学、天文、地理等专家都算



做科学家。

4. 音乐家专指：作曲、作词、指挥。

5. 单纯名人是指，该人物的出名、成就主要是由于他们与名人的某种关系，否则以相同的境遇与表现是没有被介绍的必要的，即使本人的成就很突出，仍要看是不是他们单纯名人的实质是更大的影响力量。例如：有一篇美国总统林肯回忆他的母亲的文章，中间讲到了他的母亲伟大的爱，实质上，与他母亲相似的有很多的母亲，而这位母亲由于与一位名人有这样的关系才被记载和介绍。

6. 演艺人员包括：歌唱演员、影视演员、戏曲、曲艺、舞蹈演员等以表演为主要职业的。

7. 领袖人物，专指：马、恩、列、斯、毛、邓、孙中山、周恩来。

8. 思想家、哲学家，中间包括了中国的大儒，各哲学流派的思想家等。

第五章 实 验 法

实验方法主要用于研究现象之间的因果关系。实验可以在实验室内进行，也可以在实地进行。在实验室实验中，实验对象和一切必要的材料均被置于一个被人为安排好的、可以由研究者细心控制的环境之中。实地实验则在实验室以外进行，条件比较自然，可以是在街道上、监狱、医院或工厂里进行。如果研究者想对每个微小细节都加以控制，采用实验室实验更合适。而当研究者想尽量减少人们在实验室条件下改变其行为的可能性时，用实地实验更合适。在精确新闻报道中，采用实验法多是实地实验。

菲利普·迈耶在《新精确新闻学》（New Precision Journalism）一书中，认为编辑们都很喜欢实地实验。当迈耶还是一名年轻记者时，一位聪明并且多少有些疯狂的编辑比尔·菲利普斯（Bill Phillips）派他到大街上去，做出各种粗鲁的行为，目的是为了检验迈阿密人对粗鲁行为的忍受情况，迈耶撞向急冲冲赶路的商人、把烟圈喷向一位慈祥的老妇人的脸、故意在公共汽车站的车行道上撒了一把零钱，坚持捡起每一分钱后才让汽车通行，以致交通堵塞。实验结果显示，公众对种种粗鲁行为的反应是友好和有礼貌的。

在4月1日愚人节那一天，比尔·菲利普斯和菲利普·迈耶把一个钱包故意放在人行道上，然后躲在角落里，用一个细绳把钱包拴起来，如果路过的行人试图把钱包捡起来的话，菲利普·迈耶就可以用绳子把钱包拉过来，以此来戏弄行人，实验的目的是看一下被捉弄的人有什么反应。有个摄影师躲在附近来记录被捉弄的行人的反应，实验结果显示，菲利普·迈耶的鼻子挨了不少重拳。

下面所列的是关于实地实验的一篇报道：

"丢失"钱包 考验诚实

美国一家著名杂志社最近做了一项实验，他们先后在美国、欧洲、亚洲、加拿大、澳大利亚、新西兰和拉丁美洲的一些城市，故意在人行道上、电话亭里、办公楼旁、打折商店、教堂、停车场和饭店里"丢失"了1100个钱包。每个钱包里装有相当于50美元的当地钞票，另外还装有一张印有电话号码的名片，以便捡到钱包的人可以方便地找到失主。

调查结果表明，44%的钱包被人捡走后不归还。不过，各国的情况不大一样。最为诚实的国家是挪威和丹麦，"丢失"的钱包全部物归原主。从整体上讲，斯堪的纳维亚国家的诚实程度超过了世界其他国家。新西兰、韩国和日本的情况也比较好。在美国，10个"丢失"的钱包有7个物归原主。这使得美国和加拿大、澳大利亚、印度一道并列在"较为诚实"的国家名单中。但是，如果你在阿根廷和意大利的话，可得注意保管好你的钱包。如果你在墨西哥丢了钱包，十有八九是找不回来了。当然，同一国家的不同城市区别也很大，例如，西雅图就远远好于亚特兰大。

通过在各个地方与归还钱包的人交谈，实验人员发现，虽

然世界各国具有不同的文化传统，但凡是拾金不昧者，都有一些共同因素在起作用。

一是家长对孩子的言传身教。在德国的魏玛，10个“丢失”的钱包8个不见去向。可是，有一个钱包是一个8岁的小女孩在她妈妈的指点下归还的。小女孩的妈妈说：“我小的时候，家里很困难，但父母非常诚实。我一直在用同样的方式教育我的孩子。”

二是对信仰的忠诚。马来西亚一个20岁的妇女，家里很穷，但还是归还了钱包。她说：“作为一个穆斯林，我知道存在诱惑，并知道怎么去克服诱惑。”当然，不归还钱包的人也有信教的。在墨西哥，有两个看上去很虔诚的基督教徒，在捡钱包前，还做了一个十字祷告。他们认为这是天赐之物，非拿不可。

三是将心换心。有很多次，看起来很穷的人常常归还钱包，看上去很富的人捡起钱包就一去不回。一个为躲避科索沃战乱而在瑞士当餐厅招待员的阿尔巴尼亚人，在归还钱包时说：“我每天都要干很长时间的活儿，我知道要挣这么多钱必须付出怎样的劳动。”

（摘自《环球时报》2001年9月12日第十版）

第一节　实验法的基本概念

在本节中，我们先介绍一下在实验法中涉及到一些概念，包括自变量与因变量、事前测量与事后测量、实验组与对照组、随机分配等，然后介绍进行实验的步骤，以及霍桑实验和双盲实验，最后，我们介绍一下实验法有哪些优点和缺点，以及道德方面的考虑。



一、自变量与因变量

自变量（Treatment or independent variable），也被称为处理变量，是实验中的刺激因素，可以由实验者控制的。

因变量（Dependent variable），也叫结果变量，在实验中，因变量可能是对应着不同处理的物理状态、社会行为、态度、感觉以及信念等等。因变量可以通过调查问卷、观察、访问以及生理上的反应（如心跳、血压等）来测量。

在菲利普·迈耶的粗鲁行为实验中，菲利普·迈耶的粗鲁行为就是自变量，人们对菲利普·迈耶的粗鲁行为的各种反应就是因变量。

有学者做过这样的一个实验，在校园里，坐在轮椅上的妇女依次向40名男性和40名女性求助，请他们帮忙寻找掉在林荫道上的耳环。在这个实验中，研究对象（Subjects）是80名被拦截的男性和女性路过者，自变量是伤残妇女的请求，因变量就是每个研究对象帮忙寻找耳环时所花费的时间，时间的长短是以分钟来计量的，是由躲在旁边椅子上假装看书的人观察记录下来的。

二、事前测量与事后测量

事前测量（Pretest），也叫前测，是指自变量发生以前所测量得到的因变量，即对实验对象还没有进行处理之前所测得的因变量的大小。

事后测量（Posttest），也叫后测，是指在实验状态中，引入自变量后，即对实验对象进行处理之后，所测的因变量。

在一项有关促销活动对超市的巧克力销售量影响的实验研究中，自变量是促销活动，因变量就是每个超市的巧克力的销

售量。在促销活动开始前，先对选中的超市进行事前测量，即记录下来每个超市的巧克力的销售量，促销活动开始，再记录下来每个超市的巧克力的销售量，就属于事后测量。

三、实验组与对照组

实验中常常把研究对象分为几个组，目的是进行比较。常见的是分为两个组，其中的一组接受自变量的刺激，叫做实验组（Experimental group）。实验组也叫做处理组。在实验中，没有接受处理的那一组叫做控制组（Control group），也叫做对照组。

在社会科学研究中，对照组的存在，不仅可以防止实验本身的影响，还可以排除实验进行过程中外在事件的影响。

在一项有关促销活动对超市的巧克力销售量影响的实验研究中，共抽取了20家规模、人流量、销售额、周边环境等都非常类似的超市作为实验对象，这20家超市都出售A品牌的巧克力。选择其中的10家进行促销活动，这10家超市就是实验组，另外10家超市没有进行促销活动，就是对照组。

四、随机分配

对于每个研究对象来说，是到实验组，还是到对照组，严格说来，应该是随机分配（Random assignment）的，即每个研究对象到实验组的机会与到对照组的机会是相同的。

在促销活动对超市的巧克力销售量影响的实验研究中，共抽取了20家规模、人流量、销售额、周边环境等都非常类似的超市作为实验对象，其中的10家作为实验组，进行促销活动，另外的10家作为对照组，不进行促销活动。那么，20家超市中，到底哪10家作为实验组，哪10家作为处理组，应该



是随机分配的，即每一家超市到实验组与到处理组的机会都是相同的。

五、进行实验的步骤

研究者根据研究目的，确定研究题目，并把它变为可测量的问题，用变量提出假设，用实验来检验假设。以下所列的实验步骤是一个非常理想的实验室实验程序[①]，在实际进行实验过程中，由于研究目的的不同以及各种条件的限制，有时会省略一些步骤。

1.	确定一个与实验研究相适应的假设。 Begin with a straightforward hypothesis that is appropriate for experimental research.
2.	确定一个实验设计，能操作的前提下可以用来检验研究假设。 Decide on an experimental design that will test the hypothesis within practical limitations.
3.	决定如何引进处理，或者说创造条件引进自变量。 Decide how to introduce the treatment or create a situation that induces the independent variable.
4.	完善因变量的测量的有效性和可信性。 Develop a valid and reliable measure of the dependent variable.
5.	建立实验框架，对处理和因变量的测量进行探索性尝试。 Set up an experimental setting and conduct a pilot test of the treatment and dependent variable measures.
6.	确定相应的研究对象。 Locate appropriate subjects or cases.

① W. Lawrence Neuman《Social Research Methods-Qualitative and Quantitative Approaches》，第三版，第183页。

7.	把研究对象随机分组，并给出适当的介绍。 Randomly assign subjects to groups (if random assignment is used in the chosen research design) and give careful instruction.
8.	对所有的组进行因变量的事前测量。 Gather data for the pretest measure of the dependent variable for all groups (if a pretest is used in the chosen design).
9.	只对实验组进行处理。 Introduce the treatment to the experimental group only (or to relevant groups if there are multiple experimental groups) and monitor all groups.
10.	对所有的组进行因变量的事后测量。 Gather data for posttest measure of the dependent variable.
11.	向研究对象报告研究的真正目的和实验的原因，询问研究对象的实际感受。特别是如果在实验中对研究对象的某些方面有所欺骗时，如实报告是十分必要的。 Debrief the subjects by informing them of the true purpose and reasons for the experiment. Ask subjects what they thought was occurring. Debriefing is crucial when subjects have been deceived about some aspect of the experiment.
12.	检验所收集的数据，用统计数字和图表来确定，研究假设是否被支持。 Examine data collected and make comparisons between different groups. Where appropriate, use statistics and graphs to determine whether or not the hypotheses is supported.

六、霍桑效应

霍桑效应（Hawthorne effect）是美国心理学会在20世纪30年代，于美国西部电气公司的霍桑工厂进行一项管理实验时发现的。当时，厂方迫切想要提高生产效率，想找出可以鼓励工人提高产量的方法。研究者梅奥把一组女工与其他工人分开，并有意改变了照明条件、工间喝咖啡的休息时间、吃午饭的时间和支付报酬的方法。起初，梅奥及其同事很高兴：每个变化



都带来了劳动生产率的提高。但是，研究发现，无论改变什么条件，生产率都提高，他们产生了怀疑。后来，把工作条件恢复到老样子，可工人们的劳动生产率还在提高。原因是作为实验对象的女工由于知道了自己是实验的对象，为自己所受到的注意感到高兴，她们感觉到研究者期望得到的结果，从而使自己的生产行为一反常态，劳动积极性大涨，尽最大努力来让研究者感到高兴。

在实验中，研究者要尽量避免霍桑效应，即由于实验对象对研究者试图证明的内容做出了设想而破坏了实验。

七、双盲实验

双盲实验（double-blind experiment），是指在实验中，实验处理对实验对象和与实验对象接触的工作人员来说都是未知的。研究对象不知道研究的目的，不知道自己是在实验组还是对照组，而与研究对象直接接触的研究人员也不知道这一点。实验处理是由实验对象和与实验对象接触的工作人员以外的第三者来实施的。

八、实验法的优点

实验方法之所以有用，是由于实验使研究者可以了解一些在日常条件下往往无法系统研究的特殊问题。因为在日常条件下，许多其他影响会掩盖或歪曲有关的过程。

九、实验法的缺点

实验法也有一些问题存在，它只能用来研究限定得很窄的问题。人们在人为造成的实验环境中的行为可能与他们在这个环境以外的行为大不相同。

十、道德方面的考虑

社会科学中，实验法的研究对象是人，所以，无论是在实验室实验中，还是在实地实验中，一定要考虑所进行的实验所带来的负面影响是什么。下文是一篇关于实验研究的道德问题的报道：

美国教授闹翻240家餐馆

美国著名哥伦比亚大学的教授弗兰克·弗林上个月给纽约市240家餐馆分别发信一封，抱怨自己与妻子在结婚周年纪念日进食这些餐馆的食物后出现了“长时间恶心、呕吐、腹泻和腹部绞痛”症状。

《纽约时报》昨天的报道说，哥伦比亚大学商学院教授弗兰克·弗林发出此信的意图并不是恶作剧，而是“学术研究”，希望能够考察餐饮业在接到顾客投诉后的反应。信写在了带有校名和校徽的信纸上，而研究项目其实并没有获得校方授权。

弗林发出的信件，挑明自己和妻子“严重食物中毒”。虽然是千篇一律的内容，没有提供就餐日期和时间以及选定的菜式等基本信息，但却在多数餐馆内着实造成了一场混乱，较之寻常受到客户服务投诉的情形更加让人摸不着头脑，也令餐馆老板和厨师们更加惊恐万状。

伙计们忙着在众多信用卡收费记录和外卖记录中寻找弗林的名字，希望能够确定他是什么时候“光顾”了自己的店门或者预订了自己的服务。他们还急于检查菜单上列出的各种菜式，希望能够发现到底是什么原因引起了食物中毒。谈起此事，一家法式餐馆的老板让·克洛德·贝克告诉记者，他为此把自己手下的厨师们骂了个“狗血喷头”。他承认，“那是一段神

经极度紧张的经历”。

不过，弗林的投诉信，语焉不详，也让少数餐馆老板产生了怀疑，其中就包括一家烧烤店的合伙人劳伦斯·克雷奇默。按照他的话说，“来信既没有写明就餐日期，也没有指出可能导致食物中毒的菜式……有太多疑点表明其中有诈”。

就此，在餐馆给弗林的诸多回信中，有60封对弗林的“食物中毒”投诉提出了疑问，以至于暗示投诉者存有诈骗目的的可能性。在哥伦比亚大学商学院，院长迈耶·费尔德伯格读了其中一封回信，随后向全部240家餐馆发出了一封正式道歉信。费尔德伯格认为，这是“一名低资历教员犯下的一个异乎寻常的判断错误”，而教员自作主张实施的研究项目本身也有着拙劣的构想。

前任院长干预此事之后，弗林改用普通信纸写下了一封道歉信，发往众餐馆。他解释说，先前的去信“是编造的结果，目的是为了帮助研究项目收集数据，以了解商家对消费者投诉的反应”。此事并未了结。费尔德伯格说，校方正在对这一事件进行全面回顾和彻底调查，弗林在哥伦比亚大学的职业前景还不明朗。而在弗林那里，《纽约时报》记者打到他办公室和家里的电话都无人接听，对于记者留在这两处录音电话上的留言也没有回复。

摘自《北京晚报》2001年9月9月国际新闻版

第二节 实验设计

一、经典实验设计

经典实验设计（Classical Experimental Design），也称为古典

实验设计，是最基本、最为标准的实验设计。经典实验设计包含了实验设计的全部要素：实验组、控制组、事前测量、事后测量、实验处理（自变量）、因变量以及随机分配。

具体的操作步骤是：

1．随机地把接受实验的人分成两组，实验组和控制组；

2．分别对实验组和控制组进行事前测量；

3．只对实验组进行实验处理；

4．对实验组和控制组进行事后测量。

表 5-1

	随机分组	事前测量	处理	事后测量
实验组	有	○	×	○
对照组		○		○

假如研究者的目的是了解黑人、白人同班学习如何影响白人学生对黑人的态度，最理想的方式就是采用经典的实验设计的方式，即：

1．把白人学生随机分成两组，两组学生在各个有关的方面都很相似；

2．测验两组学生在种族问题上的态度；

3．实施处理，即实验组，同黑人学生同班；另一组，即对照组，除了不与黑人学生同班外，其他经历均与实验组相同；

4．一段时间以后，再次测验两组学生的种族态度，其中任何的差异都被认为是自变量造成的结果。

我们再来看一个有数据的例子，在 1992 年的美国总统选



举中，在总统候选人电视辩论前后，以大学生为研究对象，考察了对候选人的人格魅力评价得分，满分为7分（详细资料请见本章的应用实例）。自变量为是否收看了总统候选人的电视辩论，因变量为对独立候选人皮尔特的人格魅力评分以及对共和党候选人克林顿的人格魅力评价得分。见表5-2：

表5-2

	人格魅力评价得分	事前测量（电视辩论前）	处理（是否收看电视辩论节目?）	事后测量（电视辩论后）
皮尔特	实验组	3.9	收看	5.7
	对照组	3.7	没看	4.1
克林顿	实验组	5.7	收看	5.8
	对照组	5.9	没看	5.7

从上表可知：

对皮尔特而言，实验组事后、事前测量的差异 = 5.7 − 3.9 = 1.8

控制组事后、事前测量的差异 = 4.1 − 3.7 = 0.4

电视辩论对皮尔特人格魅力得分的影响 = 1.8 − 0.4 = 1.4

对克林顿而言，实验组事后、事前测量的差异 = 5.8 − 5.7 = 0.1

控制组事后、事前测量的差异 = 5.7 − 5.9 = −0.2

电视辩论对克林顿人格魅力得分的影响 =

$$(-0.2)-0.1=-0.3$$

二、前实验设计

在实际工作中，由于各方面条件的限制，有时无法进行经典的实验设计，特别是在实地实验中。前实验设计（Preexperimental Designs）经常被非专业的、不用于推断严格的因果关系的研究所采用。

1. 单组事后设计

单组事后设计（one-group posttest-only design），也叫做一次性个案研究设计（One-Shot Case Study Design）。只有实验组，在接受实验处理之后，进行事后测量。

表 5-3

	随机分组	事前测量	处理	事后测量
实验组	无		×	○
对照组				

下面我们来看一个由菲利普·迈耶主持的关于药剂师对艾滋病认识的实地实验：

1988年，北卡洛莱纳大学礼拜山校区（University of North Carolina at Chapel Hill）医学部的研究人员 Inge Corless 博士在准备一门有关艾滋病的课程，作为研究的一部分，她与本地的一些药剂师来讨论不同类型的避孕套预防艾滋病的特性，发现这些专业人员这方面的知识是相当缺乏的。为了验证 Inge Corless 博士的发现是否适合于大部分的药剂师，在菲利普·迈耶所主持的高级报道课程上，设计了一项实验，安排学生到所在镇及



附近的另一个镇的每一个药店去购买避孕套，并询问如何有效地预防艾滋病。

所定义地区的所有药店都被调查了，不存在抽样的问题。每个药店只抽取了一名店员作为实验对象，用这个药剂师来代表这个药店中的所有药剂师。

研究者设计了一个量表来测量每个药剂师对艾滋病预防知识知道多少。这个量表主要的内容如下：

（1）橡胶比动物皮要好（后者的小孔可能让病毒通过）。

（2）润滑的橡胶比干橡胶要好（破裂的可能性要小）。

（3）用精胺（spermicide）来润滑比普通润滑要好（精胺可以杀死艾滋病病毒）。

（4）可以杀死艾滋病病毒的精胺在实验室中的学名是Nonoxynol-9。

实验结果显示，只有40%的药剂师知道所有这些简单的事实。更糟的是，一些药剂师建议学生购买者购买羊皮纸避孕套，这种避孕套价格上来说是最贵的，但在预防艾滋病的效能方面却是最低的。

这个实验是简单并且容易概括的，不需要探究因果关系和相关关系。简单的百分比本身就有足够的新闻价值。学生们试图检验一些关于无知的一些假设，比较了连锁药店和独立药店、年长的药剂师和年轻的药剂师的答案，既没有差异发现，也没有新闻价值。

这个实验结果发表于《Durham Morning Herald》，产生的社会效应有两个方面，一方面促进了本地的药剂师更好地进行学习，另一方面，直接向读者进行了有关艾滋病预防的教育。

2. 单组事前事后设计

单组事前事后设计（One-Group Pretest-Posttest Design）只

有一个实验组，对实验组进行事前测量和事后测量，没有对照组存在，所以最大的弊病是自变量以外的某些因素可能会影响因变量的变化。

表 5-4

	随机分组	事前测量	处理	事后测量
实验组	无	○	×	○
对照组				

3．静态组比较

静态组比较（Static Group Comparison）有实验组与对照组，但没有事前测量。

表 5-5

	随机分组	事前测量	处理	事后测量
实验组	无		×	○
对照组				○

我们来看一个实例：

美国北卡洛莱纳州的法律规定，18 岁以下的未成年人不准饮酒，酒吧等场所也不能卖酒给未成年人。实验的目的是检验礼拜山地区警方对非法卖酒给未成年人处罚的效应。研究第一个假设是这样的：当警方处罚时，售酒者就不敢卖酒给未成年人。为了检验这个假设，附近地区可以卖酒给未成年人的啤酒销售商作为控制组。



第二个假设是：警方一般在周末的晚上查得比较严，这样的话，工作日的晚上违法行为就会比较频繁。

在礼拜山地区，访问了所有的便利店和酒馆，在对照组中，是利用黄皮电话号码簿进行抽取的，类型、数量与实验组是相同的。

为了不违法，所安排的去买酒的学生都在18岁以上。测量变量是看啤酒的售卖者是否要求买酒者出示相应的身份证明。

共走访了246家售酒商，在星期六的晚上，礼拜山地区的销售商要求出示身份证明的有22%，而在对照组中，只有6%的销售者要求购酒者出示身份证明；单独看来，22%的比例也不是很高，似乎警方的处罚没有什么作用，但与6%相比，会发现实验组与对照组有着显著的差异。警方的处罚有显著的威慑作用。

但是到了星期一的晚上，礼拜山的售酒商要求购买者出示身份证明的只有7%了，与对照组的差异不显著了，这表明，警方处罚的威慑作用是有选择性的，只在周末明显。

表5-6

	事前	处理	事后（周六晚上）	事后（周一晚上）
实验组		处罚	22%	7%
对照组			6%	6%

三、准实验设计

准实验设计（Quasi-Experimental Designs）包括6种具体的类型，有些类型，如拉丁方设计和因子设计，比较复杂，多用于实验室实验，在实地的实验使用不多，由于篇幅所限，只给出简单的列表说明，详细的内容请参考有关的书籍。

1. 有对照事后设计

有对照事后设计（Two-group posttest only），随机分组，没有事前测量，对实验组实施处理，对实验组与控制组进行事后测量。有对照事后设计与静态组比较的差异在于对研究对象有随机分组；与经典的实验设计的差别在于实验组与对照组都没事前测量。

表5-7

	随机分组	事前测量	处理	事后测量
实验组	有		×	○
对照组				○

2. 间断的时间序列设计

在间断的时间序列设计（Interrupted Times Series）中，只有实验组，在实验处理之前，进行多次事前测量，同样，在实验处理之后，也进行多次事后测量。



表 5-8

	随机分组	事前测量	处理	事后测量
实验组	无	○○○○	×	○○○
对照组				

在美国，1979 年以前，烟草的消费税一直保持不变，1979 年烟草的消费税提高了 35%，然后 10 年内保持不变。研究假设是烟草消费税的提高会降低烟草的消费量。有学者考察了 1970 年到 1990 年间的情况，发现，在新税率实施前的 9 年，烟草的消费量保持不变，1979 年新税率实施，消费量下降，然后一直保持稳定。

3. 等时间序列设计

等时间序列设计（Equivalent time series），同间断的时间序列设计一样，只有实验组。对实验对象先进行测量，然后实施实验处理，再测量、再处理，再测量，依次类推。

表 5-9

	随机分组	测量	处理	测量	处理	测量	处理	测量
实验组	无	○	×	○	×	○	×	○
对照组								

在美国，1975 年以前，没有要求骑摩托车的人要戴头盔，1975 年颁布法规，要求骑摩托车的人要戴头盔，1981 年，在摩托车俱乐部的压力下，此法规废除，1989 年此法规恢复。

研究者的假设是戴头盔可以减少事故中摩托车手的死亡率。研究结果发现，在1975年前，摩托车手的死亡率很高；1975年至1981年，死亡率大幅下降；1981年至1989年之间，死亡率高过1975年以前的水平；1989年至今，死亡率下降。

4．拉丁方设计

如果研究中涉及到几个实验处理应该如何安排的话，就要采用拉丁方设计（Latin-Square Designs）。表5-10所列的是3种实验处理的情况：

表5-10

	随机分组	测量	处理	测量	处理	测量	处理	测量
第一组		○	×a	○	×b	○	×c	○
第二组		○	×b	○	×a	○	×c	○
第三组	有	○	×c	○	×b	○	×a	○
第四组		○	×a	○	×c	○	×b	○
第五组		○	×b	○	×c	○	×a	○
第六组		○	×c	○	×a	○	×b	○

5．所罗门四组设计

所罗门四组设计（Solomon four-group design）是最理想的实验设计，它的前两组是一个经典的实验设计，后两组是一个有对照的事后设计。



表 5-11

态度得分	随机分组	事前测量	处理	事后测量
第一组	有	○	×	○
第二组		○		○
第三组			×	○
第四组				○

表 5-12 是关于学生对老年人的态度测量，实验处理是观看一部有关老年人的经历、困难和贡献的电影：[①]

表 5-12

态度得分	随机分组	事前测量	处理	事后测量
第一组	有	85	×	45
第二组		85		75
第三组			×	60
第四组				83

实验组总的影响 = 第一组的后测 − 第一组的前测

= 45 − 85 = −40

事前测量的影响 = 第二组的后测 − 第二组的前测

= 75 − 85 = −10

以第一组和第二组前测的平均值作为第三组、第四组的事

① 数据选自风笑天《社会学研究方法》，中国人民大学出版社，2001 年 3 月第 302 ~ 304 页。

前测量的值

实验处理的影响 = 第三组的后测 - 第三组的前测

= 60 - 85 = -25

交互作用的影响 = 实验组总的影响 - （事前测量的影响 + 实验处理的影响）

= -40 - （-10 - 25） = -5

外部因素的影响 = 第四组的后测 - 第四组的前测

= 83 - 85 = -2

则：

实验处理的影响 = -25 - （-2） = -23

事前测量的影响 = -10 - （-2） = -8

交互作用的影响 = -5 - （-2） = -3

6．因子设计

因子设计（Factorial Design）用于考察两个以上自变量的同时作用。表5-13所列的是两个子变量的情形：

表5-13

态度得分	随机分组	事前测量	处理	事后测量
第一组	有	X1	Z1	○
第二组		X1	Z2	○
第三组		X2	Z1	○
第四组		X2	Z2	○



第三节 内部效度和外部效度

在实验方法中，内部效度和外部效度是关心的焦点。有些书上也称内部无效度和外部无效度，即要采取措施，避免内部无效度和外部无效度。

一、内部效度

内部效度（internal validity）是指实验结果正确反映实验本身的能力，也就是说，因变量的变化完全取决于自变量（处理），而不受其他因素的影响。在实验中，其他变量的存在使得处理变量不是因变量发生变化的真正原因，如果自变量以外的其他因素对因变量产生影响的话，将会降低实验的内部效度。

内部效度的逻辑是通过控制实验条件以及通过实验设计来排除处理变量以外的变量（影响因素）。一般说来，可能降低内部效度的因素有以下几个方面：

1. 选择性偏差

选择性偏差（Selection Bias）是指研究对象被分到每一组的可能性不同，是由于没有采用随机分配所造成的。如果实验组中的研究对象具有某种影响因变量的特征，那么，就会发生选择性偏差。如果想了解体育课对大学生身体素质的影响，实验组的学生都来自体育系，而对照组的学生都来自文艺系，那么，很显然，两组的起点是不同的，存在着选择性偏差。

2. 历史

历史因素（History）是指在实验进行中，一个与处理变量无关的事件发生了，并且影响了因变量。历史效应多发生在历

时比较长的实验中。

3. 成熟

成熟（Maturation）是指与处理变量无关的被研究对象的心理、生理、情感的变化对因变量的影响。在持续时间很长的实验中，成熟的影响是很常见的。比如，在一项有关儿童玩耍方式的实验中，对小学一到六年级的学生进行实验，随着年龄的增加，儿童的生理、情感以及心智的成熟对玩耍方式都会产生影响，而不仅仅是处理变量的影响。

事前测量和控制组可以帮助研究人员确定历史因素和成熟因素是否存在，因为实验组与控制组共同经历了时间的变化。

4. 测验

有时候，事前测量（Testing）本身会对实验产生影响。所罗门四组设计可以帮助研究者发觉测验效应的存在。比如说，在课程的第一天的，研究者让学生们进行了一次测验，课程就是处理，在课程的最后一天又进行了一次同样的测验，来检验学习的效果。如果，学生记住了事前的考题的话，就可能会影响到他们的学习，或者说影响到对事后测验的回答，测验效应就存在了。测验效应的存在，使得处理效应不是影响因变量的惟一因素。

5. 测量工具

测量工具（Instrumentation）影响稳定性信度。发生在实验过程中仪器或者因变量的测量发生变化。比如说，在减肥的实验中，体重计不灵敏了，所测的体重低于实际的体重了；又比如，在态度的测量中，所用的量表不完全相同。

6. 死亡率

死亡率（Mortality）是指一些研究对象没有完成整个实验过程，中途就退出了。在一项减肥实验中，开始时有50个人，



在实验完成时，只剩下30人了，每个人都减轻了5斤。那中途退出的20个人可能与坚持下来的30人不同，可能有人减掉了25斤后退出的，也可能减肥过程对研究对象造成了伤害，不得不退出了。研究者应该报告每一组事前、事后的人数。

7．统计的回归性

统计的回归性（Statistical Regression）是指即使没有实验处理，一些极端的研究对象也会向全体研究对象的平均值靠近。

例如，为了检验某种新的教学方法，把学生作为研究对象。成绩已经很好的学生，成绩再提高的可能性不大了，而个别成绩非常差的学生，不管有没有新的教学方法，成绩都很容易提高。在实验中，常常把好学生的成绩没有提高归结于新教法的无效，或者把差学生的成绩提高归为新教法的有效性。在实验研究中，要特别注意这一点。

8．处理的扩散或污染

处理的扩散（Diffusion of Treatment or Contamination）是指不同组的研究对象相互交流，从而知道了其他组的处理方式。在一项新的单词记忆法的实验中，持续时间为一天，实验组的成员学习了新的单词记忆法，在中间休息时，告诉了在控制组的人，从而影响了控制组成员的记忆效果。可以把实验组、控制组相互隔离，或者请求研究对象发誓保密。

9．补偿性行为

补偿性行为（Compensatory Behavior）在一些实验中，给予实验组特殊的待遇，而控制组没有。这种不同的待遇被知晓了。这种不平等可能会对控制组产生压力，更加努力以减小不同之处。各种类型的补偿性行为都可能会对因变量产生影响。例如，一个学校接受实验处理，增加中午的休息时间，来看对学习成绩的影响。这个不平等的待遇被控制组知晓后，他们要

求被同等对待，努力学习，以消除不平等。

在医学实验中，护士知道病人中谁在实验组，谁在处理组，她知道对照组的病人吃的药是安慰剂，可能会觉得对不起对照组的病人，这样在护理上就更加精心了，不同于实验组了。

10．研究人员的期望

研究人员也可能影响内部效度，不是通过不道德的有意的行为，而是通过与研究对象的非直接沟通，让研究对象感受到了研究者的期望（Experimenter Expectancy）。研究人员可能对研究假设持一种肯定的态度，期望得到能支持假设的结果。比如说，一位研究人们对伤残反应的研究者可能坚信女性对伤残的反应比男性更加敏感。通过眼神、语音、语调以及一些非语言的交流，研究人员无意中鼓励女性研究对象对伤残报告肯定的感觉，而对男性则刚好相反。

为减少研究人员的期望对内部效度的影响，一般采取双盲实验。比如说，为了研究一种新药的疗效，研究者用三种颜色的胶囊来装药，黄色的胶囊装的是新药、粉色的胶囊装的是旧药，而绿色的胶囊装的是安慰剂（也叫宽心丸，没有任何疗效，无毒无副作用）。给病人发药的人员不清楚各种颜色的胶囊所装的是什么药，这样可以如实记录病人的反应。病人不知道自己吃的是什么药，可如实地报告自己的感觉。

二、是否有影响内部效度的因素存在

下面我们通过具体的例子，来看一下，如果通过实验结果，来判断是否有影响内部效度的因素存在。

有 5 种减肥方法，分别用甲、乙、丙、丁、戊来表示，对于 5 种方法，都设计了一个经典的实验设计，表 5-14 列出了 5



种方法的实验组与控制组的事前和事后体重[①]。我们通过数据来分析一下哪种减肥方法最有效。

表 5-14

平均体重（人数）	所用符号说明			(甲)	
	事前测量	事后测量		事前测量	事后测量
实验组	A(A)	C(C)	实验组	190(30)	140(29)
对照组	B(B)	D(D)	对照组	189(30)	189(30)
	(乙)			(丙)	
	事前测量	事后测量		事前测量	事后测量
实验组	190(30)	141(19)	实验组	160(30)	152(29)
对照组	189(30)	189(28)	对照组	191(29)	189(29)
	(丁)			(戊)	
	事前测量	事后测量		事前测量	事后测量
实验组	190(30)	188(29)	实验组	190(30)	158(30)
对照组	192(29)	189(28)	对照组	191(29)	159(28)
A、B、C、D 表示平均体重，单位为磅；(A)、(B)、(C)、(D)表示人数；采用随机分配的方式确定是到实验组还是控制组。					

下表对于每一种减肥方法，列出了事前、事后的比较以及实验组与对照组的比较。其中：

① W. Lawrence Neuman《SOCIAL RESEARCH METHODS-Qualitative and Quantitative Approaches》，第三版，第 197 页。

第一列为5种减肥方法的符号；

第二列A－B表示减肥方法开始前，实验组与对照组的平均体重之差，用于判断两组的起点是否相同。如果不同，有可能发生选择性偏差；

第三列C－D表示减肥方法结束后，实验组与对照组的平均体重之差，用于判断两组的结果是否相同。如果不同，处理可能是无效的，或者可能有强历史性、成长性、扩散因素对减肥效果产生了影响；第四列C－A表示实验组的事后、事前的平均体重之差，用于判断实验组是否发生变化。如果没有变化，处理可能是无效的；

第五列D－B表示对照组的事后、事前的平均体重之差，用于判断对照组是否发生变化。如果发生变化，可能存在了影响内部效度的因素；

第六列（C）－（A）与第七列（D）－（B）表示实验组以及控制组的参加人数有没有变化。如果人数大量减少，则意味着对内部有效性有致命的伤害。见表5-15：

表5-15

	A－B	C－D	C－A	D－B	(C)－(A)	(D)－(B)
(甲)	1	－49	－50	0	－1	0
(乙)	1	－48	－49	0	－11	0
(丙)	－31	－37	－8	－2	－1	0
(丁)	－2	－1	－2	－3	－1	－1
(戊)	－1	－1	－32	－32	0	－1

根据以上结果，可以得出以下结论：



1．对于减肥方法甲：实验组的人平均体重减少了 50 磅，而在对照组的 30 人体重没有任何变化。只有 1 人在中途退出。没有明显影响内部效度的因素，处理效应显著，即减肥措施有效。

2．对于减肥方法乙：实验组的人平均体重减少了 49 磅，而在对照组的 30 人体重没有任何变化。但实验组中有 11 人在中途退出。存在死亡率过高的问题，对内部效度有影响。

3．对于减肥方法丙：实验组平均体重减少了 8 磅，而控制组平均体重减少了 2 磅。存在的问题是实验组事前的平均体重比控制组少 31 磅。有选择性偏差。

4．对于减肥方法丁：实验组平均体重减轻 2 磅，而对照组平均体重减轻 3 磅，处理效应不明显，即减肥措施无效。

5．对于减肥方法戊：实验组与控制组的体重都减少了 32 磅，意味着成熟效应、历史效应以及处理扩散效应发生。

三、外部效度和实地实验

外部效度是指实验的结果推广到实验以外的能力。即使实验结果正确地反映了实验过程，那么，实验结果能否推及到社会生活的更广阔的方面呢？我们先来看一个例子：

在社学领域，曾分别进行过实验室实验和实地实验来检验“匿名性增加侵犯行为”这个假设。实验室实验是在 1970 年进行的，让参加实验的学生，负责给另一名学生实施电击，表面的理由是“了解人们对其他人痛苦的反应”。其中的一组实验对象套上只露出眼睛的布罩，并且在黑暗中向另一人实施电击，因此，电击的实施者绝不会被认出来，是匿名的；另一组人在实验中佩带有表明身份的名牌。两组人可自主决定电击时电压的大小。实验结果表明，匿名者给予的电压远大于不匿名

者所给的电压。这一实验室实验的结果证实了“匿名性将增加侵犯行为”。

为了验证在实验室中得到的结论，在自然的环境中是否也是成立的，又设计一项实地实验。研究者根据不同社区环境中人们的匿名程度不同，来变化实验的刺激。在大城市里，人口众多，相互之间来往不多，因而人们很难被别人认出；而在小镇上，人口较少，人们之间接触比较多，很容易被别人认出来。如果研究假设是成立的话，在大城市更容易做出侵犯行为。研究者在纽约大学（匿名性强）和斯坦福大学（匿名性差）附近的街道上各停放了一量相似的汽车，26 小时后，纽约街头的汽车被卸走了蓄电池、水箱、雨刷、一个备用轮胎、一盒上光蜡以及 4 个轮盖等物品，而斯坦福大学附近的汽车完好无损。这个实地实验说明了：越是不为人所知，侵犯行为的频率就越高。

在这个例子中，实地实验的结论与实验室实验的结论是相吻合的，说明该实验的外部效度比较高。而有些实验的外部效度比较低。

有学者认为，影响实验室实验外部效度的因素主要有 4 个方面[①]：

1. *测量的反应或交互效应*

测量的反应或交互效应（Reactive or Interaction Effect of Testing），是指事前测量可能影响实验对象的对实验处理的回答或反应，所说的影响即可能是增加也可能是减少。

比如，在被研究对象观看一部有关化工业对环境污染的影

① Natalie L. Sproull《Handbook of research methods-A Guide for Practitioners and Students in the Social Sciences》，第二版，第 140 页。



响的影片之前，如果研究对象先回答了一份有60个题的反对化工的问卷，那么，事前填答这份问卷，可能会使影片的影响增加。

为避免测量的反应或交互效应因素对外部效度的影响，建议不进行事前测量。

2．选样和实验变量的交互作用

如果在选样过程中有偏差存在的话，就可能产生选样和实验变量的交互作用（Interaction of Selection and Experimental Variable)。

有研究者想在一家公司内进行一项实验，所联系的前11家公司都拒绝了请求，而第12家公司十分愿意合作。很显然，在是否接受实验方面，这家公司与前11家公司有着显著的差异，这样的话，第12家公司中的研究对象可能会对实验处理更加接受或更加敏感。

避免选样和实验变量的交互影响的办法是，增大实验对象的代表性，即在每家公司只抽取较少数量的研究对象，但是增加抽取的公司的数目。

3．实验分配的反应效应

实验分配的反应效应（Reactive Effect of Experimental Arrangements)，是指实验对象可能以某些未知的方式对实验条件做出反应。

在一项对系主任的研究中，系主任们意识到自己成为了研究对象，会去猜想实验的目的，并做出相应的反应，就是说，这时候研究对象的反应不仅仅是针对实验处理的反应了。

避免实验分配的反应效应的方法是不进行事前测量，并且不让研究对象意识到有实验在进行。

4. 多重处理的冲突

多重处理的冲突（Multiple-Treatment Interference），是指如果针对同一研究对象实施了多重处理的话，前面的处理会对后面的反应有影响，从而无法分辨出每一种处理的影响。

一组财会人员先参加一项如何与审计人员合作的培训，接着又接受了如何与经理合作的培训，因为培训的影响不能消除，先进行的培训可能会影响后进行的培训。

应尽量避免出现针对相同的研究对象实施多重处理。

应用实例

一、研究背景

1992年在美国总统选举期间，美国Connecticut大学教授祝建华博士和J.Ronald Milavsky博士，利用总统候选人在电视上公开辩论的机会，进行了有关接触传媒对知识获取以及印象评价的影响效果的实验研究。当时，3位总统候选人分别是乔治·布什、比尔·克林顿和罗斯·皮尔特，其中，布什是现任总统，民主党候选人；克林顿是共和党候选人，胜出希望很大；皮尔特是独立候选人，知名度不及布什和克林顿。

3位总统候选人将在1992年10月11日进行第一次电视公开辩论。

实验的研究对象是185名Connecticut大学的选修传播课程和演讲课程的学生，参与实验是课程的一部分，学生将获得学分。研究对象中有58%是女生，1/3是一年级学生，1/3是二年级学生，1/3是三、四年级学生。一半左右的学生有兼职或全职的工作。



二、实验设计和实施

1. 事前测量：3 位总统候选人乔治·布什、比尔·克林顿、罗斯·皮尔特第一次电视公开辩论是在 1992 年 10 月 11 日。在 10 月 7～9 日，对 185 名大学生的全体进行了第一次调查（事前测试），通过自填问卷了解他们对 3 位总统候选人的印象评价，以及对候选人关于各种问题所持观点的了解程度；

2. 随机分组：从 185 名学生中随机地抽取 53 人作为实验组；将其余 132 人中曾在公开辩论当晚观看过电视辩论的 65 人作为自然观看组；再将余下的 67 人随机地分成读报组（46 人）和控制组（21 人）。共计分成 4 个组；

3. 实验实施以及事后测量：在 10 月 11 日电视辩论当天，实验组的 53 名学生集中在一个指定的地点，观看了总统候选人电视辩论的实况转播，观看完毕后填答调查问卷，内容与事前测试的相同；

10 月 12 日，首先了解实验组以外的 132 名学生是否观看了昨晚的电视辩论，将已经在家里观看了电视辩论的 65 名学生归入自然观看组，请他们填答调查问卷，内容也与事前测试的相同；

将没有观看电视辩论的另外 67 人随机地分成读报组（46 人）和控制组（21 人）。安排读报组先阅读当天的《纽约时报》（The New York Times）对总统候选人电视辩论的报道，然后填答相同的调查问卷；控制组的学生则安排直接填答调查问卷。

实验设计列表汇总说明见表 5-16：

表 5-16

组别	名称	人数	是否随机分组	事前测量	处理	事前测量
第一组	实验组	53	是	有	集中看电视	有
第二组	读报纸组	46	是	有	看报纸	有
第三组	控制组	21	是	有	没有接触媒体	有
第四组	自然观看组	65	否	有	在家看电视	有

三、部分实验结果说明

自变量和因变量：实验中的自变量用变量 GROUP 表示，取值分别为 1、2、3、4。不同取值意味着在实验中接受的处理不同。具体如上表所示。

在这项研究中，包括很多个因变量，总的说来包括两个大的方面：

一方面是一些客观题，是测量对总统候选人的所持的政见的了解程度，所持的政见包括对国内、国际等等事务。

另一方面是对总统候选人的印象评价，由一个印象评价量表构成。印象评价的量表参考了 MILLER ET AL 的五维评价（a five-dimensional scheme developed by Miller et al）体系，即从 5 个方面对候选人进行评价。每个方面用 3 个具体的指标/变量来表示，即整个印象评价量表一共有 15 个变量。每个变量采用了 7 级测量的尺度，“7” 表示最肯定的回答，“1” 表示最否定的回答，6、5、4、3、2 的评价等级介于 7 和 1 之间。

15 个具体的测量指标详见表 5-17，说明如下：

第一列表示 15 个指标在问卷中出现的先后次序，在问卷



中，有关观点的知晓程度和印象评价的问题是穿插在一起的；

第二列给出了变量名称，从中可以看出 Miller et al 的 5 个维度和各个变量之间的关系，即 COM1、COM2、COM3 表示“Competence（能干）”这个维度，CSK1、CSK2、CSK3 表示“Communication skills（沟通技巧）”，INT1、INT2、INT3 表示“Integrity（正直）”，CHA1、CHA2、CHA3 表示“Charisma（个人魅力）”，POT1、POT2、POT3 表示“Potency（权势）”；

第三列给出了 15 个指标的中文意义；

第四列表示每个指标的“7”的含义；

第五列给出了测量的等级，7 级量表；

第六列表示每个指标的“1”的含义。

表 5-17

出现次序	变量名称	中文意义	“7”的含义		“1”的含义
1	COM1	有才智	Intelligent	7 6 5 4 3 2 1	Not Intelligent
2	COM2	能干的	Competent	7 6 5 4 3 2 1	Incompetent
3	COM3	知识广博	He knows a lot	7 6 5 4 3 2 1	He doesn't know much
4	CSK1	善于表达	Expressive	7 6 5 4 3 2 1	Unexpressive
5	CSK2	条理清楚	Clear	7 6 5 4 3 2 1	Vague
6	CSK3	明确的	Speaks in specifics	7 6 5 4 3 2 1	Speaks in generalities
7	INT1	值得信任	Trustworthy	7 6 5 4 3 2 1	Untrustworthy
8	INT2	真诚的	Sincere	7 6 5 4 3 2 1	Insincere
9	CHA1	举止文雅	Pleasant	7 6 5 4 3 2 1	Unpleasant
10	CHA2	亲切的	Warm personality	7 6 5 4 3 2 1	Cold personality
11	POT1	敢作敢为	Aggressive	7 6 5 4 3 2 1	Passive

续表

出现次序	变量名称	中文意义	"7" 的含义		"1" 的含义
12	POT3	强壮的	Strong	7 6 5 4 3 2 1	Weak
13	POT3	支配的	Dominant	7 6 5 4 3 2 1	Subservient
14	CHA3	友好的	Friendly	7 6 5 4 3 2 1	Unfriendly
15	INT3	诚实的	Honest	7 6 5 4 3 2 1	Dishonest

下面我们来看一下，实验研究中不同组别的研究对象对3位总统候选人的个人魅力评价得分。"人格魅力"变量是一个总加变量，由3个具体的评价人格魅力的7级指标总加后取平均而成，即：

$$CHA = (CHA1 + CHA2 + CHA3)/3$$

其中，CHA表示"人格魅力"，CHA1表示"举止文雅（Pleasant)"，CHA2表示"亲切（Warm)"，CHA3表示"友好(Friendly)"，CHA1、CHA2、CHA3的取值范围都是从1到7，数值越大，表示越肯定。

结合分组均值以及方差分析的结果，我们可知：

布什的人格魅力：事前测量表明，不同组别的大学生对布什的人格魅力评价得分非常接近，结合方差分析，我们知道，差异不显著（方差分析的显著性水平为.437)；事后测量显示，电视辩论以后，不同组别的大学生对布什人格魅力的评价也没有显著差异（方差分析的显著性水平为.901)。

克林顿的人格魅力：事前测量表明，不同组别的大学生对辩论前的评价没有显著差异（显著性水平为.451)；事后测量显示，电视辩论以后，不同组别的大学生对辩论后克林顿的人格魅力的评价也没有显著差异（显著性水平为.974)。



皮尔特的人格魅力：事前测量表明，各个组别对辩论前的评价没有显著的差异（显著性水平为 .738），事后测量显示，电视辩论以后，不同组别的人对辩论后的皮尔特人格魅力评价是有显著差异的（显著性水平为 .000），实验组（观看电视）、自然观察组和读报组对皮尔特的人格魅力的评价得分，在电视辩论后有显著的提高，而控制组的辩论前后的评价差异不大，见表 5-18：

表 5-18

人格魅力得分均值	布什		克林顿		皮尔特	
	辩论前	辩论后	辩论前	辩论后	辩论前	辩论后
实验组	4.5	4.6	5.7	5.8	3.9	5.7
读报组	4.7	4.7	5.7	5.7	4.1	4.9
控制组	4.9	4.7	5.9	5.7	3.7	4.1
自然观看组	4.9	4.8	5.9	5.7	3.9	5.2

综合方差分析的结果和以上的评价得分均值情况，我们可以得出以下结论：

第一，对于像布什、克林顿这样公众比较熟悉的人物，电视辩论并没有增加对其人格魅力评价的得分；

第二，对于皮尔特这样知名度不太高的人物，电视辩论可以帮助观众加深对其的了解，可以提高对其人格魅力的评价得分。

第六章 调 查 法

调查研究（survey research）方法是精确新闻报道中最常使用的收集数据的方法。调查研究是指采用问卷的方式，直接地、系统地从一个样本中收集资料，并利用统计分析方法对资料进行处理。

在调查研究方法中，涉及抽样，即如何选择调查对象，这一方面的内容请见第三章；在调查中，有一份围绕研究目的所设计的调查问卷，问卷质量的好坏对调查由直接的影响，有关问卷设计的内容请参见第八章；问卷调查得到的是大量的数据，要利用统计软件进行统计分析，这部分的内容请见第九、十、十一章。

在本章中，我们重点介绍调查方法中的数据收集过程，即访问实施过程。

在调查研究方法中，按照是由访问员进行提问还是由被访者自己填答问卷，可以分为访问式调查和自填式调查。访问式调查是指访问员按照问卷的问题顺序，向被访者提问，然后把被访者的回答记录在问卷上，被访者本人是看不到问卷上的问题的；自填式调查，是指由被访者自己来填写问卷，填写的时候，访问员可以不在旁边，访问员即使在旁边，也只是起到答



疑的作用。

按照数据收集方式的不同，调查访问主要可以分为3种：面访调查（face-to-face interview survey）、电话访问（telephone survey）和邮寄问卷调查（mail survey）。下面我们来详细介绍。

第一节 面访调查

面访调查，顾名思义是访问员与被访者相见，面对面地进行访问；面访调查按照访问地点的不同，可以分为入户面访和街头拦截面访。

一、入户面访调查

入户面访，是指接受过培训的访问员按照抽样方案中的要求，到抽中的居民家庭中去访问。

1. 入户面访的过程

(1) 找到居民户

一般说来，入户面访中，哪个居民户作为调查对象，在抽样方案中都有所说明，不需要访问员自己确定，多数时候在抽样方案中都具体地给出了待访问居民户的具体地址，访问员的任务是找到相应的地址。一般说来，访问员不得随意更换被访居民户。

(2) 说服居民接受访问

找到相应的居民户之后，访问员要向居民说明来意，请求接受访问。近年来，由于社会经济的发展，个人隐私的观念也进一步加强，越来越多的人不愿意陌生人来打扰自己的生活；另外，有些地方，由于社会治安不好，人们的安全防范意识很强，很多人家都安装了防盗门，一见是陌生人来访，很多人都

不会开门，更不愿意接受调查。拒访的原因可能是没有时间、对访问内容不感兴趣、认为侵犯隐私、对调查主办机构不信任等等。有时，拒访者不愿意说明拒访的原因。绝大多数的调查都属于自愿性调查，即被访者没有配合调查的义务。所以，对于拒访，要予以理解。

一般说来，能否说服居民接受访问，一定程度上取决于访问员的访问技巧以及对调查项目的介绍。为了取得居民户的信任，访问员要带好自己的证件以及调查机构的证明。

(3) 抽取被访者

在访问员的努力下，一旦抽中的居民户同意接受访问，那么，就要确定是由家里的哪一个人来接受访问。在抽样方案中，一般会给出入户后，如何抽取调查对象。

(4) 实施访问

在实施访问的过程中，如果是访问式问卷，访问员要与被访者面对面坐好，访问员要严格地按照调查问卷的指示以及培训的要求来进行提问，有些时候，要出示一些辅助的卡片或资料，被访者根据自己的情况对所问问题给予回答，访问员要及时、准确地记录下每个问题的答案，对开放式的问答题一般要进行充分的追问。

如果是自填式问卷，访问员可以把问卷交给被访者自己来填写，对于被访者有疑问或者是不清楚的地方，访问员要按照事前的培训内容给予解答。

(5) 检查问卷并致谢

在问卷中所有的问题都问完或填完以后，访问员要从头到尾检查一下问卷，以免遗漏问题。对于有疑问的地方，要向被访者明确地提出来，及时加以解决。

在确认访问完成以后，要向被访者表示感谢。在商业性调



查中，为了感谢被访者的配合，有时会在调查结束时，向被访者赠送礼品。

在入户访问时，访问员要针对抽样要求中的每一个人填写情况说明。

2．入户调查的优点

入户面访调查最大的优点可获得有代表性的样本。在入户调查的抽样时，可以保证随机的原则，如果所抽中的家庭能够配合的话，所获得的样本是具有代表性的。基于这一点，很多大规模的、要求对总体进行推断的调查都是采用入户调查的方式，如收视率调查等。

入户面访调查是在被访者所熟悉的环境里，问卷可以长一些，询问的内容可以多一些，同时可以使用辅助工具，如向被访者展示相关资料等。

3．入户调查的缺点

入户调查最大的缺点是成本高，被访者的住处距离调查机构可能比较远，用于旅途的时间和费用都要计入成本。

质量控制困难，一般说来，为降低成本，每个访问员都是单独工作的，容易出现访问员弄虚作假的现象。

入户困难，成功率低，由于人们隐私的观念越来越强，绝大多数调查又是自愿性的，被访者没有责任要接受访问，所以，访问的成功率比较低。如果访问的成功率过低的话，所获得的样本会有偏差存在。

二、街头拦截式面访调查

街头拦截式面访调查主要有两种方式：

第一种方式是由经过培训的访问员在事先选定的若干个地点，如繁华的交通路口、大的商场内（外）等，按照一定的程

序和要求，例如，每隔几分钟拦截一位，或每隔几个行人拦截一位，等等，进行抽样拦截。对所抽取的对象，向其说明介绍调查的目的，在现场按照问卷进行简短的面访调查。这种方式常用于需要快速完成的探索性研究。

第二种方式也叫中心地调查（central location test）或厅堂测试（hall test）。按照项目的要求，在繁华地段拦截可能合适的访问对象，问几个简单的问题，来判断被拦截者是否符合筛选问卷的要求，如符合条件，征得其同意后，带其到附近事前租好的专门地点，进行详细的面访调查。这种方式常用于广告效果测试、新产品品尝等需要进行实物展示的市场调查项目。

如某品牌的洗发水广告投放市场前，设计了几个广告版本，想了解一下哪个版本效果最好，就可以进行厅堂测试。目标消费者可以是年轻的女性，高收入、高文化程度。那么，在拦截时，就要找可能符合条件的人，然后把合适的、愿意配合的人带到附近的一个公共场合，在事先租好的房间里看几种版本的广告录像，然后回答有关问题。

街头拦截式面访调查的主要优点是节省费用，同时减少访问员作假的机会。

街头拦截式面访调查主要缺点是样本代表性差，所获得样本对总体不具有代表性；另外，如果访问时，访问员与被访者都是站着进行的，则问卷不宜太长。

三、电脑辅助的面访调查

在面访调查中，无论是入户面访还是街头拦截面访，都可以使用电脑来进行辅助调查，即不再使用传统的印刷问卷，而是把问卷放在电脑里。同时，不再需要问卷录入电脑的过程。利用调制解调器，访问员每天可将调查数据直接传回公司。



计算机辅助个人面访调查 CAPI（Computer-Assisted Personal Interviewing）是指访问员看着屏幕的提示来提问，及时地将答案通过键盘、鼠标或专用电脑笔输入计算机内。

电脑辅助自填式问卷 CASI（computer assisted self-interviewing）是对被访者进行简单的培训或指导后，让被访者面对电脑屏幕上的问卷，通过键盘、鼠标或专用电脑笔，逐题地将自己的答案亲自直接输入计算机内。调查员不参与回答，也不知道被访者输入的答案，但是调查员可以待在旁边，以便随时提供必要的帮助。

CAPI 以及 CASI 在一些发达国家使用比较广泛。所用的电脑有两种形式，一种是普通的笔记本电脑，另一种是专用的无键盘轻型电脑（Pen-Table PC），重量不到 2 公斤，配有可用 2 个小时以上的又小又轻的蓄电池，利用电脑笔可将被访者的回答写进电脑。

在一些发达的国家，每年用 CAPI 以及 CASI 完成的调查量已经大大地超过了传统的用纸和笔来完成的面访调查（PAPI）。例如，荷兰的 NIPO 公司在 1996 年完成了 100 万个访问，有 40%的 CATI（电脑辅助的电话调查）、50%的 CAPI 以及 10%的传统纸加笔的面访调查（PAPI）[①]。

第二节 邮寄问卷调查

邮寄问卷调查，就是事先获得被访对象的邮寄联系方式，然后将调查问卷装入信封，通过邮局或其他渠道寄给选定的调查对象，并要求被访对象在规定的时间内，按规定的要求填写

① 柯惠新等《市场调查与分析》，中国统计出版社。

问卷，然后寄回调查机构。

2000年北京广播学院团委、党办等部门对在校学生的家长进行了一次邮寄调查，目的是了解学生家长对学生在校生活的了解程度，以及对学校的期望等。

一、邮寄调查的基本步骤

一般说来，邮寄调查要经历地址的收集、事前联系、寄出问卷、催促等几个步骤。

1．收集被调查对象的通信地址

根据调查研究的目的来确定调查的总体是什么人，收集所有的调查对象的名单、通信地址或电话。这种资料收集工作常常是十分困难的，需要通过多种直接的和间接的渠道，才有可能获取一份完整的、有效的邮寄名单。

在北京广播学院对学生家长的调查中，学生家长的姓名以及通信地址都是现成的，在学生入学登记表中就有，收集起来比较容易。

如果被调查总体很大，可以按照一定的原则从中抽去一部分作为邮寄问卷的对象。

2．事前联系

在正式的调查问卷寄出之前，可以先通过电话、明信片或简短的信件，与被调查对象进行初步接触，说明最近将有一份邮寄的问卷，请求他们协助填写。

一般认为，这种事先的接触可以显著地提高邮寄问卷的回收率，但是在经费或时间不允许的情况下，这一步骤常常被省略了。

3．寄出邮件

通过邮局或其他邮递渠道向调查对象发出调查邮件。一般



说来，在寄出的邮件中要包括调查问卷、有关的说明信以及回邮信封。

(1) 调查问卷

调查问卷是所寄出邮件的核心内容，所有要问的问题都在问卷中。在设计问卷时，特别注意问卷的措辞要便于自填。

(2) 致调查对象的说明信

致调查对象的信，也可以写在调查问卷的封面。在信中一般要说明此次调查的组织机构、调查的目的、问卷的填写要求等。此信的目的是恳请被调查者合作，语气上要非常客气。一般要加盖调查机构的印章，如果项目负责人知名度比较高的话，也可以请项目负责人签字。有些商业性的调查，为了提高回收率，常常许诺被调查者如果寄回问卷的话，有一定的鼓励措施，如赠送小礼品、礼金或参加抽奖等。这些许诺常常写在信中。

2000年人民日报《新闻战线》编辑部采用邮寄调查的方式，对《新闻战线》的读者进行了调查。致调查对象的封面信如下：

读者朋友：

我们《新闻战线》杂志全体工作人员向您致意！如您所知，《新闻战线》是党中央机关报——人民日报主办的综合性新闻专业刊物，也是新中国第一份新闻业务刊物。自1957年问世以来，一直面向全国新闻工作者，包括新闻院校师生、新闻研究人员和广大新闻爱好者，为办好报纸、广播、电视特别是党报，发挥了重要作用。

为在新世纪里更好地传达党中央对新闻工作的指示、要求，引导新闻工作者学习、运用马克思主义、毛泽东思想和邓小平理论，按照党中央的指示和要求，指导新闻工作实践和新

闻理论研究，现开展一次读者随机抽样调查，请您在问卷括号中打勾，书写您的意见、建议，以便我们改进工作。

为保证调查的科学性、完整性，请您务必将填好的问卷放在随此函邮来的信封中于8月10前寄回。

谢谢您的支持和合作！

此致

敬礼

人民日报《新闻战线》编辑部

（此处加盖了公章）

我们再来看一个台湾的例子，是由台湾政治大学学者苏衡给出的[①]，介绍信是这样写的：

各位先生、女士：

你好。今年将举办中央民意代表选举，定在某月某日举行投票。我们认为，选举是大家的事，每一个人都应该有机会对这次选举表示自己的想法，所以我们从全省的选民中，用抽签的方法抽出100人，他们也跟您一样，收到这份问卷。请您仔细阅读，详细说明您自己的意见，只要10分钟左右。

既然是意见，就没有对错的问题，请您按照自己的意思放心填写。

本机构过去曾经做过多次选举研究，并曾公布研究结果，获得社会大众良好的反应。今年选举，本机构再度进行研究，希望了解全体民众对这次选举的意见。

我们也跟您一样，认为只有大家都尽量表示意见，民主政治才会进步。您提供的意见我们保证保守秘密，绝对不会把您的意见告诉任何人。我们关心的，是全体民众的整体意见，不

① 苏衡《传播研究调查法》，三民书局，1986年，第107页。



会透露任何个人的姓名以及其他资料。

请您在一周内填毕，寄还给我们。随函附上已经贴好邮票的信封，您不必贴邮票，我们希望您能尽量在一周内填毕寄回。如果您不能及时寄回问卷，我们将再去函催促您尽快填毕，寄还给我们。如果您有任何疑问，请打电话给我们，可由我们付电话费，我们的号码是××××××××。

谢谢您的合作。敬祝

身体健康事业顺利

××机构负责人×××敬启（亲笔签名）

××年×月××日

（3）回邮信封

回邮信封一般采取印刷的形式，印好了调查机构的通信地址，同时贴足了邮资。被调查者只要把填好的问卷装入信封，投入邮筒即可。如果采用自邮寄问卷（self-mailing）的话，就可以省去回邮信封。自邮寄问卷是指答卷人可以按标好的虚线把问卷折叠成一个信封，自己写好地址，贴好邮票，直接投入邮筒即可。

4．第一次催促

国外的研究发现，被访者在收到问卷后，两三天内填好、寄出得比例是最高的，随时间的增长而递减。为了提高邮寄调查的回收率，还要进行催促。

第一次催促一般在问卷寄出一周后进行。通过电话或简短的提示信，与调查对象再次接触，询问其是否收到了问卷，并请求合作早日寄回问卷。

这种再次的接触往往也可以显著地提高回收率，但是在经费或时间不允许的情况下，这一步骤也常常被省略。

5. 第二次催促

对回收的问卷及时登记编码，包括问卷的编号、问卷寄回的日期、寄回的地区，按回收的日期统计回收的数量。如果回收率还未达到理想的水平，则再次打电话或寄提示信（第二封提示信最好再次附上一份问卷）；一般在问卷寄出3周后进行第二次催促。

也有学者认为，如果第二次催促后，回收率还是不够理想的话，可以进行第三次催促。一般在问卷寄出7周后进行，同时再次寄出包括介绍信以及问卷。

在进行催促时，要强调此项研究的重要性，被访者的意见是不可替代的，提醒其马上填写问卷并寄出，并表示感谢。在措辞上要注意有礼貌，语言婉转客气，避免被访者反感。

二、邮寄调查的主要优点

1. 保密性强

由于邮寄调查一般都是匿名的，被访者可以自主地选择合适的时间来填写问卷，然后用统一的回邮信封寄回，所以保密性很强。被访者不用担心访问员会知道自己的答案，有安全感，可以完全按自己的实际情况来填写问卷。因此邮寄调查比较适合于敏感问题的研究。

2. 调查区域广

从理论上来说，凡是通邮的地方都可以调查。因此面访或电话调查难于寻找到的调查对象，例如，边远地区或海外地区的被访者，都有可能通过邮寄的方法将问卷寄去。

3. 费用较低

显然，邮寄调查的费用比面访调查低；一般情况下也比电话调查的费用低。



4．无调查员偏差

邮寄调查没有涉及到访问员，被访者有充分的时间和自由度来回答问卷中的问题，可以完全避免由于调查员的原因而产生的偏差。

三、邮寄调查的主要缺点

1．回收率低

在几种调查方法中，邮寄调查的回收率是最低的。根据美国一个研究机构的报告，如果调查对象是随机选择的，在没有任何事先和事后的接触，又没有谢礼的情况下，邮寄调查的回收率在美国一般不会超过15%，[1] 在我国可能更低（还没有专门的研究结果）。因为调查对象可能对所调查的内容不感兴趣，或是由于问卷过长过于复杂，等等。因此在邮寄调查中，如何采取有效的措施提高回收率是非常重要的。

著名的社会学研究方法专家艾尔·巴比认为，要进行分析和报告撰写，问卷的回收率至少要有50%才是足够的（adequate）；要至少达到60%的回收率才算是好的（good）；如果回收率能达到70%就非常好（very good）。[2]

当然，最重要的是，要保证所回收的问卷没有偏差。如果经过验证，一个没有偏差的样本，回收率就是低一点，也比有偏差的高回收率的样本要好。

2．时间周期长

在几种调查方法中，邮寄调查所需的时间是最长的。因此邮寄调查只适用于那些时效性要求不高的项目；对于精确新闻

① 柯惠新 丁立宏编著《市场调查与分析》，中国统计出版社。

② 艾尔·巴比《社会研究方法》（上），华夏出版社，第331页。

来讲，由于对时效性的要求比较高，很少采用邮寄调查的方式。

3．填答问卷的质量难于控制

调查对象可能会找他人代为回答，或没有填答全部问答题就停止了，也可能只挑选自己感兴趣的问题来回答，等等，这些都将影响数据的质量。

4．调查对象的限制

邮寄调查的最大限制之一是被访者必须有较高的文化程度。住在已经通邮的地区，以一般居民或消费者为总体的调查不适于用邮寄调查。有研究表明，在邮寄问卷调查中，拒访者多是低教育程度的民众。[①]

四、电脑辅助的邮寄问卷调查

电脑的发展，尤其是互联网的发展，使邮寄问卷调查进入了一个新的阶段。省去了印刷问卷的环节。电脑辅助表现在两个方面：一是自填式问卷的电脑化（computerized self-administered questionnaire，CSAQ），即被访者通过软盘、互联网公告栏或其他方式来读取问卷，填答后再返还，可以通过传统的邮寄方式，也可以通过互联网；二是利用互联网络，用电子邮箱代替了传统的邮寄地址。

随着互联网络的发展，通过电子邮件的形式来发送问卷正在成为一种潮流。

第三节　电话调查

电话调查是以电话为中介，访问员同被访对象不需要见

① 苏蘅《传播研究调查法》，三民书局，1986年，第114页。



面，而是通过电话来交流。

一、传统的电话调查

传统的电话调查使用的工具是普通的电话、普通的印刷问卷和普通的书写用笔。经过培训的调查员在有电话的地方，可能是办公室，也可以是家中，按所规定的拨号方法拨打电话。如果拨通，则按照问卷的要求，对被访对象进行提问，同时将答案记录在问卷上。

二、计算机辅助电话调查

在精确新闻中，最常使用的数据收集方式就是电脑辅助的电话调查，为此，我们专门设立一章来介绍电脑辅助的电话调查，这里就不重复了。详细介绍请见第七章。

三、电话调查的主要优点

1．节约费用

与面访调查相比，电话调查省去了差旅的费用，成本是比较低的。特别是针对本地区的调查，不需要长途话费，费用就更低了。

2．节约时间

在几种调查方法中，电话调查的速度最快。

四、电话调查的主要缺点

1．样本的代表性问题

如果希望电话调查的样本对所在地区的全体民众有代表性的话，采用电话调查的前提就是居民住宅电话的普及率要达到一定程度，一般认为要达到90%以上。

有电话的家庭与没有电话的家庭在很多方面有着本质的不同，美国的研究表明，在美国住宅电话普及率已经达到95%以上，没有安装电话的家庭平均年龄比较低，收入比较低，教育程度较低，多是黑人，家庭中没有男性成年人，多是单亲家庭。

2. 调查的内容难于深入

电话访问的时间不宜过长，问题不宜过于复杂，因此难于调查比较深入的问题。电话调查的问卷一般都比较简短，很少采用复杂的态度量表。

3. 不能展示辅助的资料

电话调查中，由于访问员与被访者没有见面，所以无法展示有关的辅助资料。

第四节　三种调查访问方式比较

从所获得样本的代表性、问卷设计、答案的真实情况以及实施管理等方面来看，传统的入户面访调查、邮寄调查和电话调查三种方法各有所长。详细的比较见表6-1[①]：

在调查实践中，可以根据研究的目的和性质，灵活地选择数据的收集方式。香港某学者要了解港人在深圳嫖娼以及包二奶的情况，派访问员到深圳罗湖关口深圳一侧，按照一定的抽样原则，对进入深圳的香港男性进行拦截，访问员询问一些初步的问题，确认是否是合适的被调查对象，如果是的话，则请被访者接听访问员所携带的手机（移动电话），在电话中由另一访问员询问有关要了解的敏感性问题。由于是拦截式的电话访问，被访者的顾虑较小，这样的话，获得真实答案的可能性

① 参考苏衡《传播研究调查法》，三民书局，1986年，第24～26页改变而成。



就更大。

表 6-1

		入户面访调查	邮寄调查	电话调查
获得有代表性的样本	总体单位是否全部列出?			
	1. 全部列出	高	高	高
	2. 未全部列出	高	中	中
	抽样控制	高	中	高
	访问到所选对象	中	高	高
	替换被访者	中	低	低
	回答率:			
	1. 异质性高的总体	高	中	高
	2. 同质性高的总体	高	高	高
	知晓拒答理由	高	低	高
问卷设计	问卷长度	长	中	中
	问题的形式:			
	1. 复杂性	高	中	低
	2. 开放式问题的回答	多	少	多
	3. 跳答题目顺序成功率	高	中	高
	4. 控制作答顺序	高	低	高
	5. 冗长问题的回答率	高	低	中
	不回答某些问题的可能性	低	中	低
	问卷结构要清晰	低	高	中

续表

		入户面访调查	邮寄调查	电话调查
获得真实答案	社会期许的影响	高	低	中
	访问员的影响	高	低	中
	受其他人的影响	中	中	低
	需要时可询问旁人的意见	中	中	低
行政管理的需要	找到合适的访问员	低	不需要	中
	时间周期	慢	慢	快
	费用：			
	1．每份问卷费用	高	低	中
	2．交通差旅费用	高	低	中
精确新闻应用		中	中	高

第五节　大众媒介在发表调查结果时应考虑的问题

对于媒介在报道民意调查结果时，应该注意什么问题，美联社执行编辑协会 APME（The Associated Press Managing Editors Association，网址为：http：//www．apme．com）、美国民意调查研究协会 AAPOR（The American Association for Public Opinion Research，网址为：http：//www．aapor．org）和美国全国民意测验评议会 NCPP（the National Council on Public Polls，网址为：http：//www．ncpp．org）都给出过建议。



NCPP 的原则共有 20 条，是由美联社/NBC 新闻调查（Associated Press/ NBC News Poll）的创始人 Sheldon R. Gawiser 和 G. Evans Witt 给出的。下面针对第一版的具体条目给出了解释：

第一，谁做的民意调查？（Who did the poll?）

首先要明确民意调查是由哪家机构完成的？是专业的市场调查公司、学术研究机构、大众媒介本身，还是政党或其他团体完成的。调查机构的声誉对于判断民意调查的真实性、准确性是非常有帮助的。

受调查结果影响的利益团体所做的民调要特别加以注意。请看一个台湾最近 6 次关于民众对“一国两制”支持率的调查结果（来自《深圳特区报》2001 年 7 月 24 日 A4 版“民调”变味成“官调儿”）的例子：

“民调”变味成“官调儿”

6 次调查的结果如下表所示：

	时间	谁做的民调	赞成“一国两制”的比例
1	6 月 5 日	台湾无线卫星新闻台	31%
2	6 月 21 日	《中国时报》	29%
3	7 月 1 日	《联合报》	33%
4	7 月 14 日	新党立委冯沪祥委托专门机构	47.5%
5	不明	台湾外交部	10%
6	7 月 17 日	台湾行政院陆委会	13.3%

比较之后，可以看出“官”（官方）“民”（民间）之间，真是泾渭分明！所有民间机构调查的支持率，都比官方的高，

至少高出一倍；民间机构调查的结论是“支持率不断升高”，官方的却是“锐减”。……在“行政院陆委会”公布调查结果的次日，就有人不客气地公开指出，“陆委会”、“外交部”的民调充满着负面提示。也就是说，这两个官方部门是采用“特殊技巧”诱导民意，得出其预先设定的“理想结论”的。

本来，为了保证民意调查的客观性，调查方向被调查民众提问时应尽量不加预设前提，但台湾“陆委会”设计的问卷却故意“技术犯规”，预设的前提竟有3个之多，即让民众在“中华民国不再存在”等3个前提下选择“是否赞成‘一国两制’”。这样的题目设计，调查结果就不能反映真正的主流民意了。

（《深圳特区报》2001年7月24日A4版）

第二，谁出钱做的调查以及为什么做这个调查？（Who paid for the poll and why was it done?）

不管是什么机构出钱做的调查，总会有它的目的，有些专业的调查机构为了扩大影响，做一些民意调查来发表，起到宣传的作用；某些商业公司为了促销自己的产品；政治人物为了了解民众的看法等等，都可能出钱来做调查。

作为大众传媒，在发表民意调查的结果时，要考虑到，出钱做调查的机构的动机可能会应影响调查的公正性，或者说，受众可能会对民意调查的正当性产生质疑，从而使媒体的名誉受损。在国外的政治选举活动中，某个候选人可能通过民意调查来检验应对某个议题持什么样的立场，在调查中可能运用了有引导性的问题，或者采用了代表性比较差的抽样，这一类的民意调查的目的并不是反应民众的意见看法，所以在发表时，一定要慎重。

第三，在调查中共访问了多少人？（How many people were interviewed for the survey?）



调查中共访问了多少人？换句话说，调查的样本量是多少？调查10个人得出的结论，与调查1000人所得出的结论的分量是不同。在其他条件相等的前提下，样本量越大，抽样误差越小，调查结果的精确程度越高。

第四，被访者是如何抽选出来的？（How were those people chosen?）

被访者被抽取的方法就是抽样方法，一般说来，民意调查的目的是反映普遍的民意，反应大多数人的意见看法。而在实践中，由于各种条件所限，只能对一部分人进行访问，那么就要求被访问的人对所要了解的民众有代表性。在“抽样”这一章中我们已经谈到，有些抽样方法的代表性强一些，而有些抽样方法的代表性差一些，或者说，就某些抽样方法而言，假设所抽取的样本对总体有代表性风险很大。

第五，被调查者来自什么地区？来自什么人群？（What area-nation，state or region-or what group-teachers，lawyers、Democratic voters，etc. -were these people chosen from?）

抽样总体是全国性的、省的或某一具体地区，以及抽样总体的特征是什么都应该明确，是教师、律师还是某一特定人群，都应该明确界定。

第六，调查结果是否建立在所有被访者的回答之上？（Are the results based on the answers of all the people interviewed?）

有时，某些利益团体出于利己的目的，只报告了一部分被访者的意见。

第七，哪些应该被访问，但却没有成功的人是些什么人？（Who should have been interviewed and was not?）

在调查中，访问的成功率是非常重要的。如果拒绝接受访问的人与接受访问的人在个人背景特征上有明显的差异的话，

则调查结果的代表性值得怀疑。

第八，民意调查执行的时间？（When was the poll done?）

访问实施的具体时间是什么？例如，调查民众对恐怖行为的认识和态度，在“9·11”美国被袭击事件之前和之后所进行的调查，相信调查结果会有很大的不同。

第九，调查访问的方式是什么？（How were the interviews conducted?）

在调查中，收集数据的具体方式是什么？是入户面访调查、电话调查还是邮寄问卷调查？

第十，调查结果是来自于打入电话、邮寄问卷或是优惠券调查吗？（Is this a dial-in poll, a mail-in poll or a subscriber coupon poll?）

这类调查结果本质上是自愿者抽样，很难确定自愿者的回答能否代表所有目标群体的回答。

在中国，许多报纸、杂志以及网络媒体经常进行这类的调查，一般说来，回答者的数量比较多，几千、几万甚至更多。这些自愿者与那些非自愿者可能在行为、态度等方面存在着显著的差异。有人把这些自愿者称为热心读者、热心观众。

但值得肯定的是，报纸、杂志的随刊问卷调查对于改进编辑工作还是有帮助的。

第十一，调查结果的抽样误差是多少？（What is the sampling error for the poll results?）

在概率抽样的抽样设计中，会给出抽样的误差是多少。比如，在收视率调查中，抽样误差是3%，那么，如果甲节目的收视率是2%，乙节目的收视率是1%的话，考虑到抽样误差水平，实际上甲节目和乙节目的收视率高低是不可分辨的。

详细的内容请见“抽样”一章。



第十二，还有哪些错误可能影响调查的结果？（What other kinds of mistakes can skew poll results?）

除了抽样误差以外，还有很多因素可能会对调查结果产生影响，即非抽样误差，如访问员的因素、被访者的因素、问卷设计不当等等。

特别要注意的是，有时所获得样本不是自加权样本，为了对总体有代表性，对样本常常要进行加权处理。

详细的内容请见有关“抽样”一章。

第十三，所问的问题是什么？（What questions were asked?）

在调查中，所问的具体问题是什么，必须明确地给出来。同样的内容，如果采用不同的问法，措辞的不同，得到的答案可能是不一样。

详细的内容请见“问卷设计”一章。

第十四，问题的提问顺序是什么？（In what order were the questions asked?）

一般说来，在一项调查中要问很多问题，问题的提问顺序不同，对调查结果会有一定的影响。

比如，在美国进行的对克林顿总统的影响和声望评价调查，如果先问到对美国经济形式的看法，然后再问对克林顿的支持程度；另一种方式是先问对克林顿性丑闻事件的了解程度，然后问对克林顿的支持程度。相信两者得到的结果是不同。

第十五，对于同样的主题，其他类似的调查结果是什么？结论是否相同？如果不同的话，为什么会有差异？（What other polls have been done on this topic? Do they say the same thing? If they are different, why are they different?）

对于一些热门的话题，可能会有多家机构进行调查，大家

的结果可能接近，也可能相差很远，要搞清楚差异的原因。访问时间的不同、访问方式的不同、问题的措辞以及问题提问的顺序都可能对调查结果产生影响。

第十六，如果民意调查说选举已经结束，现在又如何呢？(So the poll says the race is all over. What now?)

此条是针对有关选举的民意调查的，有时候，在事前的预测调查中，可能某个候选人获得了压倒多数的支持率，但这不意味着选举结果已经出来了，什么事都可能发生，在投票前的最后一分钟，有人会改变主意，原来飘浮不定的人也会拿定主意。总之，选举前的民意调查结果并不等于真正的选举结果。

第十七，民意调查是否可以成为募集基金的工具？（Was the poll part of a fund-raising effort?)

有时，有些机构以调查问卷作为幌子，目的是推销保险或某种商品，甚至是进行募捐。这样的调查也是没有意义的。

第十八，我已经问完了所有的问题，答案看起来也不错。民意调查的结果是否就是正确的呢？（So I' ve asked all the questions. The answers sound good. The poll is correct, right?)

按照概率抽样所进行的民意调查，调查结果通常是有代表性的。但由于抽样误差和非抽样误差的存在，民意调查的结果也可能与实际的真实结果存在差距。

第十九，既然可能存在着这么多的问题，我们应该报告调查的结果吗？(With all these potential problems, should we ever report poll results?)

一般说来，调查机构是想客观、公正地反映民众的意见、看法。抽样科学、实施科学的民意调查是反映民众意见的好方法。媒体当然应该进行报道。

第二十，这项民意调查有报道的价值吗？(Is this poll worth



reporting?)

对于一项设计科学的民意调查结果，判断一项民意调查是否具有新闻价值，与判断某一事件是否具有新闻价值的标准应该是一致的。

对于大众传播媒介来说，当然要报道那些具有一定价值的民意调查结果。

在第二版的《关于民意调查结果，记者应问的 20 个问题》(20 Questions A Journalist Should Ask About Poll Results)[①] 中，结合近几年出现的新情况，Sheldon R. Gawiser 和 G. Evans Witt 对 20 个问题做了一些修订，去掉了第一版中的第十、第十六和第十七条，新增加的三条分别是："网上调查的本质?"(What about polls on the Internet or World Wide Web?)、"谁在领先?"(Who's on first?)以及"强迫性民意调查是什么?"(What about push polls?)。下面我们简单解释一下：

"网上调查的本质?"互联网的发展，使得网上调查风靡一时，网上调查本质上是自愿者调查，是由参加者来决定是否回答问题，而不是由抽样人员决定谁来回答问题。网上调查与电话打入调查、报纸杂志的随刊问卷调查一样，不知道对总体是否具有代表性。

"谁在领先?"主要是指如果两个候选人的得票不相上下，差异在误差允许的范围内的话，就无法判断出谁在领先。

"强迫性民意调查的本质?"强迫性民意调查本质上是电话行销，借打电话来诱导选民，向选民灌输关于某候选人的错误资料，影响该候选人的声誉，目的是进行竞选宣传。

---

① 《关于民意调查结果，记者应问的 20 个问题》第二版的英文全文，大家可参见http://www.ncpp.org。

第七章　电脑辅助的电话调查 CATI

用于精确新闻报道的调查，对时效性的要求很高，在最短的时间内抽取样本、实施访问并获得结果是非常重要的，为达此目的，最适合的访问方式是电脑辅助的电话调查系统，CATI 系统。CATI，是 Computer Assisted Telephone Interviewing System 的简称。

1876 年贝尔发明了电话，20 世纪 60 年代，Chiltun Research 发明了电脑辅助的电话访问系统[1]。在美国，随着经济的发展，住宅电话的普及程度也越来越高，1963 年，有 80% 的美国家庭拥有住宅电话，1980 年，有 93% 的家庭拥有住宅电话，伴随着电话普及率的提高，CATI 技术在 20 世纪 70 年代以后广受重视，得到飞速发展。

最近几年，随着我国社会经济的快速发展，一方面，居民住宅电话的普及率在迅速提高，特别是在城市地区，使得电话访问正逐渐成为可能；另一方面，经济的快速发展和社会的变迁，使得人际间的信任度越来越低，个人隐私的观念逐渐加强，传统的面访调查实施起来越来越困难，也促使电话访问成

① 苏蘅《传播研究调查法》，三民书局，1986 年，第 3 页。



为发展趋势。

目前，内地有多家机构设立了 CATI 访问中心，多是一些规模比较大的，位于北京、上海或广州的专业市场调查公司，另外一些教学科研单位也设有 CATI 系统，如北京广播学院新闻传播学院、中国人民大学统计学系都设有 CATI 系统，其中北京广播学院的 CATI 系统设有 30 条访问专线。

第一节　CATI 系统的访问过程

在这一节中，我们介绍 CATI 系统的构成，包括硬件、软件要求。并以台湾辅仁大学统计系民意调查中心的有关资料为例，来介绍一下 CATI 系统的工作流程以及如何应对访问过程中可能出现的问题。

一、CATI 系统的硬件设备

计算机辅助电话调查通常是在一个装备有 CATI 系统的专门、固定地点进行。CATI 机房表面看来，与一般的计算机机房、电脑语音教室相类似，要求温度和湿度适宜，每个位置之间有隔板。有些教学单位的 CATI 机房不仅可以从事电话访问，也可以承担电脑机房的功能。

CATI 系统的硬件设备包括有：

• 一台起总控作用的计算机服务器；

• 若干台与主机相连的终端/计算机，包括鼠标（有的还包括键盘）；

• 与终端相连的耳机/耳塞式电话；

• 与终端相连的若干条电话线；

• 监视用计算机；

• 配套的音响设备；

• 按照软件的不同，有时还需要语音调制解调器（Modem）。

二、CATI 系统的软件要求

文字处理软件有 Words、WPS 等多种，统计分析软件有 SPSS、SAS 等多种，与此类似，CATI 系统相应的软件也有很多种，大部分是英文版的，在台湾和香港有中文版的，不论是哪家机构开发的，一般说来，CATI 软件都可以实现以下功能：

• 随机抽样：按照给定的原则，电脑自动进行抽取电话号码的工作；

• 自动拨号：每个访问员电脑终端的访问电话由服务器资料库产生，并同时进行拨号，访问员不用手动拨号；

• 问卷设计：问卷利用软件直接写好，每个问题按照顺序显示在屏幕上；

• 线上访问：访问员根据屏幕所指引的问卷题目进行访问，用鼠标或键盘来记录答案，完成的问卷自动进入数据库；

• 即时监控：可随时对任何一个访问员的工作进行监听；

• 即时分析：可随时对已经完成的访问进行统计分析。

台湾辅仁大学统计系谢邦昌教授主持开发的 FJS-CATI 系统是以 Windows95 为作业环境的，在台湾应用广泛，在内地也有多家机构采用此系统，表 7-1 列出了 FJS-CATI 系统所能实现的功能：[1]

---

① 《统计教育》，2001 年第 3 期，第 46 页。



表 7-1

方向	详细功能
问卷设计	1. 图形模式编写问卷 2. 对象导向题目编辑 3. 可设定同一题目选项或不同题目间选项的互斥逻辑 4. 处理单选、复选、复选跳题规则 5. 设定各种单选、复选跳题规则 6. 随机选择 7. 随机选项 8. 线上加减题目 9. 线上加减选项
样本抽取	1. 外排抽样方法 2. 主机分派调查样本 3. 样本数据更新 4. 建立最小抽样区域
访问实施	1. 计算机自动拨号 2. 免持听筒 3. 占线/无人接听定时重拨 4. 按问卷逻辑实时检查错误 5. 对单一问题作线上提示 6. 无法回答暂时跳题 7. 约访自动提示 8. 访问员成绩记录
即时分析	1. 拒访统计 2. 访谈时间统计 3. 统计资料显示 4. 输出至 SAS 格式

续表

方向	详细功能
其他	1. 主机可随时监视任何单一工作站 2. 主机可随时传送讯息到每一个工作站 3. 计算机语音拨出并接听电话按键输入 4. 可扩充回传（fax-back）功能

三、访问的过程

下面，我们分几个方面来介绍访问的流程以及在访问中如何处理可能出现的一些问题。

1. 访问的流程

各个调查机构的CATI访问流程是类似的，不同之处在于商业性机构，是受客户委托来执行项目的，而媒介下设的民意调查中心是为媒介服务的，调查项目可能是媒介自己选定的。下面所列的是台湾辅仁大学统计系民意调查中心的访问流程：

（1）开会决定调查主题

根据政治、社会、经济等现象，于正式进行调查前3天，即根据所发生的现象，决定出所要调查的主题。

但若有紧急突发状况，如某些突发性的政治事件时，则必须临时决定是否依据这个主题为调查题目进行调查。

（2）确定受访对象及样本数

根据调查主题来确定调查对象，如特定年龄层、性别等，此外，也必须依主题来决定抽样地区，如主题为“高雄市民生活满意度调查”，抽样地区就必须针对高雄市，而非其他县市。

（3）问卷设计

根据调查主题设计问卷，内容必须要能涵盖主题，并且能



配合事后分析。为配合电话访问，访问题目最好不超过 30 题（包含个人背景资料）。

(4) 问卷审核

进行问卷审核时，需每题每句每字逐步审核，若有句意不顺或题意不清，应予以修改或删除。

(5) 问卷上线

审核完毕的问卷须将其放入 CATI 计算机系统中，将其计算机化，作为访员电话访问的依据。

(6) 计算机系统进行抽样

根据受访对象的类型及所确定的样本数，由服务器进行计算机系统抽样，抽出实际样本数及替代样本数，并由计算机自行判断何时递补替代样本。

(7) 管理人员测试

访问员进行正式电话访问之前，须由管理人员进行整体测试，以确定：

◇计算机能指定电话正常拨号；

◇访员所输入的资料能正确传回服务器；

◇若受访者拒访或空号，计算机能正确使用替代样本供访问员继续调查。

若一切测试无误，则进行下一步骤。

(8) 讲解问卷

正式调查前 30 分钟由管理人员讲解问卷并进行访问员问卷浏览，使访问员熟悉问卷，以利于访问时能把题目意义正确地传递给受访者，以降低访问时所产生的误差。

(9) 访问员进行访问

访问员正式利用 CATI 系统进行访问，根据主题收集出受访者的资料，以利于分析。

（10）管理人员监看、监听

受访者进行访问时，管理人员可利用服务器端的系统进行监看监听，以了解访问员访问情形，若有错误或资料作假，可实时纠正。同时，通过督导人员的监听、监看，可以评估访问员能力以确定其适任性。

（11）实时分析及资料备份

第一，根据每日的访问样本数进行实时数据分析，以了解每日调查状况。

第二，每日资料务必进行备份，备份于可擦写式光盘（Magnetic-Optical Drive and Disk），以防止资料损毁或流失。

（12）作业完成

调查结束后进行进一步统计分析，如：独立性检验，变异数分析等，并根据分析结果提出分析报告以供各部门参考。

2．CATI 访问的时间

一般说来，CATI 访问的时间应该是居民家里的人比较齐全的时候，同时又不会影响被访者的休息。通常，约定俗成的访问时间如表 7-2 所示：

表 7-2

工作日	星期一至星期五	晚上	18:15 ~22:00
节假日	星期六 星期日 节假日	上午	9:30~12:00
		下午	13:30~17:30
		晚上	18:15~22:00

3．访问开始前的培训

在新项目开始前，由研究人员和督导人员对访问员进行培



训。讲解问卷，对可能出现的问题加以解释。

培训时，配合培训人员的说明，访问员要详细阅读屏幕上的问卷，不仅要能够清晰、顺畅地念出题目，更要对题目的逻辑、题意的精神有清楚的认识。

如果访问员觉得题目的表达比较拗口、对题意不清楚、对专门术语或特殊字句不明了等，应及时提出来，请培训人员予以解答，必要时，对问卷进行修改。

4. 拨号

经过培训的访问员戴上配有话筒耳机或耳塞（以代替普通的电话，使双手空出来），坐在电脑屏幕前，按照屏幕上指示的程序进行工作。

自动随机拨号系统会根据研究人员事先设计好的抽样方案，自动拨号并保存拨号记录；有些 CATI 系统是在屏幕上显示出所抽取的电话号码，由访问员来手工拨号。

以下情况视为电话未接通：所拨的号码在铃响 8 次无人接听；电话占线；电话为办公电话；电话是空号；电话是传真机。

5. 开场白

如果所拨打的电话是住宅电话，则视为接通电话。完整、清晰而有礼貌的开场白，对于获得受访者的合作是十分重要的。完整的开场白一般应包括几个部分：问候语；自我介绍、调查机构介绍；调查目的；合格受访者的条件。

避免用如下的寒暄用语："能不能访问你？""访问你可不可以？""我现在可以进行访问吗？""我访问你好不好？"等，这类寒暄语是疑问句的方式，需要被访问者同意，容易被拒绝。

建议采用比较直接、肯定的讲法，如"我们想跟您做个电

话访问，以下耽误您几分钟的时间”、“现在要耽误您几分钟的时间，请你做个电话访问”等，说完之后，不要停顿太久，否则易让受访者有拒访的机会。

6. 对拒访倾向的处理

有时，开场白过后，被访者有拒绝接受访问的倾向，可能会提出各种问题，表 7-3 给出了常见的问题及其解答：

表 7-3

问　　题	解　　答
“你们是哪里？要干什么？” “你们是什么单位？要调查什么？”	对调查结构和目的简单说明，不要偏重细节，“这只是在做一项电话访问，只是在做调查研究的意见收集，只要耽误几分钟的时间。”
“你们怎么知道我家的电话号码？”	“这个电话号码是从计算机中自动拨号（随机）抽出来的，我们不知道您的姓名、地址。” “访问结果只作汇总分析，关心的全体的结果，不会作个别分析，受访者的个人意见及背景资料绝对保密。”
“我对这方面问题没有意见/不清楚/不了解。” “对这个没有兴趣。”	“我们不能随便找人访问。” “在经过追问过后，真的是没有意见，您也可以回答没有意见或者是不知道，这本身就是一种答案。” “每个人的意见都很重要。”
“我太忙了。”	“我们只有几个简单的问题而已。” “最多耽误您几分钟的时间。”

续表

问　　题	解　　答
“我拒绝接受访问。”	对此情况应加以判断，因为有些拒访实际上是要求保证不透露个人隐私，要求访问员态度的诚恳； 若仍坚决拒绝，应礼貌回答：“抱歉，打扰了”，不可争论或批评。

7. 约访

如果符合条件的受访者当时不在家（外出、上班、购物等）或者不方便接受访问（如：忙、有客人、用餐中、不舒服、睡觉等），则要进行约访，即确定受访者何时回来，方便接受访问的时间是什么时候，并将约访时间输入计算机，CATI系统也会自动地储存预约访问的电话号码和时间，届时该电话号码会自动地出现在拨号系统中。

约访时，参考问法如下：

“请问他(她)什么时候会在(回来)；通常什么时候会在家?”

“如果现在不方便，我什么时候再打电话比较方便?”

“麻烦您转告他（她），我们要做个电话访问，我们将会再与他（她）联络，谢谢。”

8. 访问进行

如果被访者同意接受访问，访问员按照屏幕上显示的问题进行询问，同时利用计算机的键盘或鼠标将数据录入计算机。屏幕上每次只出现一个问答题，计算机会根据答案自动地跳到下一道相关的问答题。调查员只需根据屏幕上显示的问答题提问，计算机会自动检查答案的适当性和一致性。

对于复选题时，受访者如果只回答一项，访问员需要追问："还有吗?""还有哪些?"

如果被访者回答类似"看情形"、"很难说"的答案，访问员需要追问受访者的倾向。

如果发现受访者不理解题意，则不要解释，更不要举例，而应以比较口语化的方式表达题目，必要时重述题目。有时，受访者对访问员的询问会不知如何回答，甚至答非所问，这可能是语言出了毛病。也许是访问员题目念得太快，受访者一时反应不过来，也许访问员说得太简略，也许说得太抽象，也许是访问员的形容太情绪化，这些都会造成被访人的困扰，应极力避免。使用的语言，以能达意为原则；用字遣词则以受访者能了解为原则。

若受访者做模棱两可或做不确切的回答，应客气地再追问，请其做出肯定的回答。此外，访问员可以主动地以各种方式反问受访者："您是不是这个意思?""您的意思是……对吗?"以及"换句话说，您认为……"

若受访者中途停止发表意见，请耐心说服，并保持礼貌。无法说服时，不要争论或批评，说："抱歉，打扰了。"

若受访者批评问卷的设计，请委婉的加以说明："当初问卷的设计就考虑到要兼顾各方面的意见"、"也许我们考虑不周，希望您提供宝贵的意见，以弥补我们的不足。"并将其评语记录下来，不要争论。

若被访者对某个问题有意见，一直不正面回答问题，则可说："那您对这一题是拒答"、"如果您对这一题有意见，您可以拒答。"以及"您有拒答的权利。"

当受访者发表意见时，有些话可能是访问员不愿意听的，有些话是题外的。原则上访问员不要打断，耐心地听下去，做



一个极有耐心的好听众。选择一个最有利的机会，使对方察觉不出来，这会让受访者觉得他是很重要的，而乐意继续回答下去。

访问员在访问过程中，要严守价值中立的原则，不加入自己的意见，并且对被访者的意见不加评论。

9. 督导员现场管理

通过 CATI 主机，可以随时提供整个调查的进展情况、每一个访问员完成工作的具体情况以及阶段性的调查结果。研究人员可以根据阶段性的调查结果及时地调整方案，使调查更为有效。例如，样本中某一类被访者的数目已经达到，下一阶段的访问就需要筛选掉这类被访对象等等。

督导员可在现场检查和指导调查员的工作，也可以在其他的房间通过监控设备随时了解访问的情况，如某个访问员可能语速太快，或是对开放题没有充分地追问等，可及时提醒访问员加以注意；访问员在遇到困难时，可随时向督导员求助。

四、CATI 访问的主要优点

1. 迅速、访问周期短

利用 CATI 进行访问，最主要的优点是缩短访问周期，可以迅速获取有代表性的数据。

利用 CATI 系统软件问卷设计控制语法容易，可以很容易地设计出研究者所需要的题型、跳答方式、选项形式，从而大量减少在问卷设计时的文字处理工作。

另外，CATI 系统不需要印刷问卷，也不需要对访问员的工作进行复核，更不需要对问卷进行电脑录入，可以节省在问卷印刷、资料录入、勘误检验上的时间。

笔者访问过的台北全国意向顾问股份有限公司的电话访问

中心配备有60条电话访问专线，调查题目在10个以内，1068个有效样本可在1天内完成；题目在20个以内的，1068个有效样本可在2天内完成。在国外媒体下设的民意调查中心，如果下午有重大新闻事件发生，从晚上6点开始调查，到晚上9点就可以知道民众对此的意见、看法。

2. 费用低

采用CATI进行访问还有一个优点就是费用低。因为CATI不需要印刷问卷，不需要复核以及问卷录入。就访问实施这个环节来看，在海外，由于人力成本较高，而电话的话费比较便宜，所以CATI比面访的费用要低很多；在内地，由于人力的成本不是很高，而电话本身的成本比较高，目前来看，同样长度和难度的问卷，CATI系统电话访问的成本和面访相差不多。但从长远来看，CATI的访问成本还会进一步下降。

近年来，一个令人瞩目的发展趋势是大型的呼叫中心（Call Center或Calling Center）也开始建立CATI系统，承接调查访问的任务。呼叫中心主要从事客户服务以及电话销售工作，许多商业性公司的客户服务工作是由呼叫中心来承担的。从硬件设备来看，呼叫中心的要求更高一些，完全可以覆盖CATI系统；呼叫中心的客户服务以及电话销售业务集中在工作日的上班时间/白天，而电话访问的时间是在晚上或休息日，两者刚好可以互补。并且呼叫中心拥有庞大的通讯集成资源（包括市话、语音中继、IP、DDN等通讯手段），大大降低了拨出电话的成本。

五、CATI访问的主要缺点

1. 不能展示有关的辅助资料

在电话调查中，访问员与被访者没有面对面接触，不能展



示有关的辅助资料。比如，在广告效果调查中，需要向被访者展示广告的草稿，以征求意见。类似这样的研究，就不能采用电话调查的方式。

2. 调查的内容难以深入

电话的时间不易过长，过长会引起被访者的反感，被访者只要挂断电话就可以中止访问。所以访问的内容不易过多，也不宜询问过于难理解的问题。

3. 有一定限制条件

在美国，1963 年，有 80%的家庭拥有住宅电话，1980 年，有 93%的家庭拥有住宅电话。即使电话普及率已经这么高了，必须注意到，有电话的人与没有电话的人还有很多不同，如南部的居民没有电话的比例是其他地区的两倍，没有电话的人群中黑人的人数是白人的 3 倍，失业者中没有电话的比例是就业者的两倍①。

如果研究的内容与是否有电话无关的话，可以忽略有电话者与无电话者的差别，而采用电话调查的方式。

如果研究主题是了解与失业有关的问题，采用电话调查就值得斟酌了。

在调查报告中，一定要清楚地写明，电话调查所代表的结果只能推断到所有拥有家庭住宅电话的人，而不是所有的人。

第二节　电话调查的抽样

一般说来，CATI 系统所拨打的电话号码是指住宅电话，

---

① Sheldon R. Gawiser & G. Evans Witt《A Journalist's Guide to Public Opinion Polls》Praeger Publishers，1994，第 63 页。

在这一节中，我们来介绍一下住宅电话调查的抽样方法。住宅电话调查的抽样分为两个阶段，第一阶段，抽取住宅电话号码，第二阶段，抽取被访者。

一、电话号码的抽取

1. 利用电话号码本抽样

电话号码本抽样（sampling telephone directory）是指与电信部门进行联系，取得完整的住宅电话号码，作为抽样框，从中采用随机抽样或系统抽样的方式获得电话号码。

利用电话号码本抽样，所拨号码为空号或所打电话为非住宅电话的情形基本上不会发生。

但值得注意的是，公开发行的住宅电话号码簿是有遗漏的，即不是所有的住宅电话号码都登记在住宅电话号码簿上，当这种情况过多的时候，抽取的样本的代表性就会受到影响。有资料表明，在美国，中小城市的住宅电话号码簿上，有10%左右的住宅电话号码没有登录；在大城市，如纽约、芝加哥以及华盛顿等，未将电话号码登录在电话号码簿上的用户，占一半左右。在中国内地，虽然没有确切的数字来说明公开发行的住宅电话号码簿中刊登的比例是多少，但相信未登的比例会更高。

特别注意的是，利用电话号码本抽样的前提是得到比较完整的住宅电话号码目录，并且要定时更新。在台湾，宇音电脑辅助电话访问系统就是采用电话号码本抽样的方式，宇音系统中包含台湾地区各县市460万住宅电话的资料，是电子版的，并且定时更新。

2. 随机拨号法

随机拨号法（Random Digit Dialing）是利用随机数字来生



成电话号码。

一般说来，电话号码由 11 位数字构成，都是由区号、局号（4 位或 3 位数字）和后 4 位数字构成的，如北京、上海、广州等中心城市，区号为 3 位，局号为 4 位，后 4 位是局号下数字，北京广播学院调查统计研究所的电话是 010（区号）6577（局号）9296（后 4 位数字）；还有一些城市区号是 4 位，如长春的区号是 0431，则城市内电话号码为 7 位，局号为 3 位，后 4 位是局号下数字。

就每个城市而言，区号是惟一的、固定的，局号的数目不是很多，大多在几十到几百之间，通过与电信部门联系或查找有关资料是可以找到局号的抽样框的。局号下的 4 位数字并不都是电话号码，电信部门在发放电话号码时不会一次把所有电话号码都分配出去，通常考虑到未来的发展，会预留一定的号码；每个局号下的电话号码数是不同的，每个局号下的住宅电话的数目也是不同，即有些局号下办公电话多，住宅电话少；而另外一些局号下，则是住宅电话多，办公电话少。

随机拨号法就是利用局号的抽样框资料随机抽取局号，后 4 位的数字随机产生。很显然，这样产生的号码可能是空号，即使是电话号码，也可能是办公电话。在访问时，需要进行判断。

随机拨号法的优点是可以找到没有登录在电话号码簿上的住宅电话。

3．电话号码加一法

电话号码加一法（plus one）是电话号码本抽样和随机拨号法抽样的结合，目的是找出没在电话号码簿上登记号码的住宅电话用户。做法是先利用电话号码簿抽取所需的电话号码，然后把这些电话号码的最后一位加上一，变成一个新的电话号

码，如65779296加1变成65779297。用这种方式产生的电话号码同样会存在空号以及是办公电话的情况，但一般说来，空号的机会比直接的随机拨号法要少一些。

二、受访者的抽选

抽取的电话号码为住宅电话时，就合乎要求。一般说来，一个家庭中生活着一个以上的人，那么，应该抽取哪个人作为访问的对象了？如何抽取呢？目前，常见的方法有5种，即是不加选择法、男女轮流选择法、楚德—卡特排列法、海根—柯林修正法以及最近过生日者法。每种方法都各有利弊。

1. 不加选择法

不加选择法（No-selection Method）就是只要接听电话的人符合调查的要求，那么谁接听电话就访问谁。比如在对总统选举的预测调查中，被访对象是选民，年龄在18岁以上，如果接电话者的年龄大于18岁，则访问此人；如果是家中的小孩子接听的电话，请他（她）让家中的大人来接听电话。

不加选择法的最大优点就是操作方便。在台湾的民意调查中，经常采用这种方法，并且有关的研究也证明了，用这种方法所抽取的样本在个人特征等方面的构成与总体是一致的。但国外也有研究发现，当电话铃声响起时，女性接听电话的比例通常是男性的两倍，即每打3个电话，有两次是女性接听的。另外，年轻人接听电话的比例也高于老年人。如果是这样的话，样本的代表性就比较差，所抽取的样本不能看成是自加权样本。

在内地，由于电话调查开始的比较晚，有关调查抽样的方法的实证研究还比较少，对于居民家庭接听电话的习惯还缺乏了解。如果从操作方便的角度考虑，采用了不加选择法来抽取被访者，那么，为保证样本的代表性，应该根据有关的总体资

料，对样本的进行加权处理。

2．男女轮流选择法

男女轮流选择法（Male / Female Alternation）是指对每个访问员来说，如果第一个被访者是男性的话，则第二个被访者为女性，然后再是男性，依次类推。

男女轮流选择法也可以根据电话号码的尾数是奇数还是偶数来决定是访问男性还是女性。比如，如果电话号码的尾数为单数的话，则访问女性，如果电话号码的尾数为偶数的话，则访问男性。

有关男女轮流选择法所获样本的代表性如何，还有待于进一步研究证实。

3．楚德—卡特排列法

楚德—卡特排列法（Troldahl & Carter Tables）要求访问员轮流使用以下 4 个表格来选择受访者。在电话接通以后，访问员要询问第一个接电话的人两个问题，搞清楚家中年龄在 18 岁以上的人数和男性的人数，然后按照楚德—卡特排列表的指示，确定谁是被访者。4 张表轮流使用。

表 7-4

—		家中 18 岁以上人数			
		1 人	2 人	3 人	4 人以上
家中男性人数	无	女性	最年长的女性	最年轻的女性	最年轻的女性
	1 人	男性	男性	男性	最年长的女性
	2 人	/	最年长的男性	最年轻的男性	最年轻的男性
	3 人	/	/	最年轻的男性	最年长的男性
	4 人以上	/	/	/	最年长的男性

表 7-5

二		家中 18 岁以上人数			
		1 人	2 人	3 人	4 人以上
家中男性人数	无	女性	最年轻的女性	最年轻的女性	最年长的女性
	1 人	男性	男性	最年长的女性	男性
	2 人	/	最年长的男性	女性	最年长的女性
	3 人	/	/	最年轻的男性	最年长的女性
	4 人以上	/	/	/	最年长的男性

表 7-6

三		家中 18 岁以上人数			
		1 人	2 人	3 人	4 人以上
家中男性人数	无	女性	最年轻的女性	最年长的女性	最年长的女性
	1 人	男性	女性	男性	最年轻的女性
	2 人	/	最年轻的男性	最年长的男性	最年长的男性
	3 人	/	/	最年长的男性	最年轻的男性
	4 人以上	/	/	/	最年轻的男性

表 7-7

四		家中 18 岁以上人数			
		1 人	2 人	3 人	4 人以上
家中男性人数	无	女性	最年长的女性	最年长的女性	最年长的女性
	1 人	男性	女性	最年轻的女性	男性
	2 人	/	最年轻的男性	女性	最年轻的女性
	3 人	/	/	最年长的男性	最年轻的女性
	4 人以上	/	/	/	最年轻的男性

比如，当访问某一家时，刚好应该使用表 7-6，这一家有



4 口人，全部 18 岁以上，其中男性有两人，则参照表 7-6，应该访问这家最年长的男性。

楚德—卡特排列方法接通电话后，要先询问第一个接听电话的人两个问题，然后访问特定的人，不仅访问时间增长，而且多数时候要换人来接听电话，容易引起被访者反感而挂断电话，使得访问的成功率偏低。

4．海根—柯林修正法

由于楚德—卡特排列法操作起来不够简便，很多学者对之提出过修正方法，其中海根（Hagen）与柯林（Collin）所提出的方法较为简单易行。

海根—柯林修正法是把家中的合格受访者分为 4 类：最年轻的女性、最年长的女性以及最年轻的男性、最年长的男性。访问时，轮流访问这 4 种人，不必先问任何问题。

海根—柯林修正法的优点是可以节省访问的时间。但相关研究发现，按海根—柯林修正法所抽出的男性比例偏低，而且无法抽到家中年龄居中的人；另外，受访者可能对“最年长”、“最年轻”的意义感到困惑。还需要进一步地加以解释，使访问时间增加。

5．下次过生日者

下次过生日者（next-birthday method）法就是先询问家中马上要过生日的人是谁，然后请这个人来接听电话。也可以问家中谁刚刚过完生日或“家中谁的生日离访问日期最近?”

有大量的研究表明，下次过生日者法是一种有效的抽样方法，所获得样本代表性比较强。

这种方法的缺点是接听电话的人可能不知道家里每个人的生日，或者记忆错误；另外，换人接听电话也容易造成挂断电话的情况发生。

应用实例

以下方案节选自北京广播学院调查统计研究所为CNNIC第七次网民情况调查所设计的一个全国性的CATI访问方案，该方案已被CNNIC采用，在2000年底实施了调查访问，在2001年1月公布了调查结果，得到社会各方面的认可。

“网民知多少?”——中国互联网络信息中心全国调查抽样方案设计[①]

肖　明　柯惠新

一、前言

随着IT技术和互联网的飞速发展，中国的网民数量、网民特征以及上网行为成为社会关注的热点，这些信息不仅对于政府管理部门制定有关政策极具参考价值，对于各个网站以及网上广告主也有一定的商业价值。1997年，经国务院信息办研究决定，由中国互联网信息中心（CNNIC）来实施这方面的调查统计工作。CNNIC于1997年、1998年、1999年和2000年共发布了6次“中国互联网络发展状况统计报告”，对我国上网人群的数量、分布、基本特征等进行了估计，对我国网民的上网行为、习惯以及对有关热点问题的看法和态度进行了概括性的说明。

这6次调查采用是网上联机调查的方式，即在网上发布问卷，由网民自由填写，CNNIC利用技术手段去掉重复的问卷，

① 此文发表于《现场统计与管理》杂志2001年第五期。



并从中抽样。利用这种网上联机方式来调查网民的情况，引起了一些争议。

为了使调查结果更具说服力，CNNIC 决定第七次调查时，有关网民情况的内容采用传统的网下抽样调查的方式进行，希望样本能够代表全国和各省的情况。为此特委托北京广播学院调查统计研究所（SSI）设计全国调查的抽样方案。

由于 CNNIC 已经形成每半年发布一次结果的惯例，所以用于调查设计和实施的时间十分有限；同时 CNNIC 属于非盈利部门，所能承担的调查实施费用也很有限。如何在有限的费用内，尽快地得到能代表各省和全国的调查结果，是方案设计的难点所在。

二、抽样设计的基本思路

1. 总体的界定

本研究的主要目的是估计全国网民的总量、网民的构成以及上网行为。考虑到上网是有一定年龄限制的，年龄太小的孩子不可能独立上网，因此将目标总体界定为"全国所有 6 岁以上的居民"。一般说来，对于这样的全国性抽样调查，采取传统的随机入户面访的方式是比较适合的，但 CNNIC 无法承受入户面访所需的时间和费用；同时，目前上网是新生事物，网民的人数还比较少，大规模的入户面访调查也不符合效率原则。

为了实施的方便，把目标总体进一步分为 3 个子总体："没有住宅电话的 6 岁以上的居民"、"拥有住宅电话的 6 岁以上的居民（不包括住校的大学生）"以及"住校的高等院校的学生"。上网是一种对文化程度和经济条件都有一定要求的行为，首先可能是在文化程度较高、家庭经济情况较好的人群中

开始流行的，这部分人群绝大多数是拥有住宅电话的；另外，大学生也是上网的先行者，大学生中上网的比例非常高；而“没有住宅电话的6岁以上的居民”中也是有人上网的，但人数可能很少，对最终结果的影响不会很大，从效率原则和操作性原则来考虑，对于这个子总体，我们不得不忽略了。

也就是说，我们把抽样总体分为两个子总体：一个是“拥有住宅电话的6岁以上的居民（不包括住校的大学生）”，以下简称“住宅电话居民”，对抽中样本采用计算机辅助的电话调查（CATI）方式进行访问；另一个是“住校的高等院校的学生”，以下简称“大学生”，对抽中样本采用面访调查的方式收集数据。

2．抽样的总思路

对于“住宅电话居民”子总体，抽样方法为分层、PPS与等概率相结合的4阶段混合型抽样：即在每个省内，第一阶段：用PPS法抽取地区（此处的地区包括地级市和地区行署，每个地区下都包含城镇和乡村）；第二阶段：抽取电话局号；第三阶段：抽取电话号码；第四阶段：抽取调查对象。所获得的样本可以近似看成是自加权样本，对全省有代表性。

对于“大学生”子总体，采用的是分层、PPS与等概率相结合的三阶段混合型抽样。即第一阶段抽学校，第二阶段抽班级，第三阶段抽住校学生。抽取的是一个代表全国的近似自加权样本，按各省大学生在各层的分配情况，可以近似推断各省的情况。

对两个子样本的结果加权处理后，可估计各省的网民人数以及上网行为；对各省结果加权处理后，可以推断全国的情况。



3．样本量的确定

上网是一个新生事物，在“住宅电话居民”中，上网人数中所占的比例还很小，为了保证目标比例估计值的精度，采用最大允许相对误差的公式来估计上网比例最低的省所需的最低样本量。

根据 CNNIC 的《第六次中国互联网络发展状况统计报告》可以近似地估算出各省的网民人数；另外从电信部门获得了各省的住宅电话数目，参考《中国统计摘要 2000》，近似估算得各省住宅电话所覆盖的 6 岁以上的人口数，两者相除，得到了 31 个省中，“住宅电话居民”中的网民比例，其中超过 5%的只有 7 个省，不足 2%的有 7 个省，网民比例最少的为 1.5%。为保证各省的精度，取 1.5%作为估计的目标比例。

同时考虑到访问实施的费用和时间，在置信度为 95%的情况下，取最大允许相对误差为 50%，设计效应假定为 2，则各省的样本量近似为 2020。为了便于各省之间进行比较，每个省都取相同的样本量。

在“大学生”总体中，网民的比例比较高，样本量按最大允许的绝对误差公式计算，综合考虑抽样设计的精度以及 CNNIC 所能承受的调查费用和调查时间，取置信度为 95%，最大允许绝对误差为 1.8%，设计效应为 2，则总样本量为 6000。共抽取 120 所大学，每校抽取 5 个班，每班调查 10 人。

4．抽样指标的确定

在“住宅电话居民”总体中，各阶段的抽样单元、抽样指标和抽样方法如表 7-8 所示。其中第一阶段的抽样指标各地区“住宅电话数目”，是利用省一级的数据建立了回归模型，预测得到的。

表 7-8

阶段	抽样单元	抽样指标	抽样方法
第一阶段	地区	住宅电话数目	PPS
第二阶段	局号	不需要抽样指标	普查
第三阶段	电话号码	后 4 位数随机产生	随机抽取
第四阶段	调查对象	不需要抽样指标	方便抽样

在“大学生”总体中，各阶段的抽样单元、抽样指标和抽样方法如表 7-9 所示。其中第一阶段的抽样指标“普通本专科学生与研究生人数”，在抽样框中是明确给出的。

表 7-9

阶段	抽样单元	抽样指标	抽样方法
第一阶段	学校	普通本专科学生与研究生人数	PPS
第二阶段	班	不需要抽样指标	随机抽取
第三阶段	住校大学生	不需要抽样指标	等距抽取

5. 关于抽样框的说明

对于“住宅电话居民”总体，采用的是多重抽样框，大部分的资料来自于统计年鉴，有关省级住宅电话数目和抽中地区的局号资料，来自电信部门；地区一级的住宅电话数目是用回归模型预测的。

对于“大学生”总体，第一阶段的抽样框是教育部发展规划司编的《中国普通高等学校名录 98 版》（高等教育出版社），



收录了截止 1998 年 9 月按国家规定的审批程序批准设置的 1022 所普通高校。包括 15 项指标，其中有 21 所学校多项指标缺失，最终的抽样框共包括 1001 所大学。近年来，由于高教改革，有些学校进行了合并，学校名称虽然改变了、学校归属改变了，但原学校的位置、学生的管理，以及原学校的规模和性质都没有大的变化。由于本研究的目的是了解住校大学生的上网情况，所以这些变化对本研究的影响不大。

第二阶段、第三阶段的抽样框都是在访问实施过程中建立的完整抽样框。

三、"住宅电话居民"子总体的抽样

第一阶段：抽取地区

从每个省中抽取五个地区，抽样指标为地区的住宅电话数目。

1. 利用回归模型预测各地区住宅电话数目

由于所获得的住宅电话的资料仅到省级，没有地区级的资料，本研究利用省一级的经济、人口指标作为自变量，"住宅电话数目"作为因变量，通过 SPSS 软件对 31 个省的资料进行多次测算，建立了一个回归预测模型。

住宅电话数目 = 325018 + 0.02684（地区人口数）+ 0.03775（国民生产总值）+ 0.1433（社会消费品零售总额）

该模型的决定系数为 0.944，回归模型的方差分析的显著性水平为 0.000，3 个自变量的与因变量的线性联系检验的显著性水平都小于 5%，虽然 3 个自变量相关显著，但研究的目标是为了预测，追求的是决定系数最大，因此可以认为所得的回归预测模型是比较理想的。当然，利用省一级的数据做出的预测模型不一定完全适用于地区，但是在没有其他办法取得实

际的各地区住宅电话数目的情况下，采用预测模型也是不得已。

按照此回归预测模型，用各个地区的人口数、1998年地区国民生产总值以及1998年社会消费品零售总额得到了各个地区的住宅电话数目。

2．抽取地区

根据所估计的各地区“住宅电话的数目”的大小，按照PPS抽样法，使每个地区被抽中的概率，等于该地区“住宅电话的数目”与该省“住宅电话的数目”之比。

3．样本量分配

在地区多于5个的省中，各省的样本量（2020）在抽中各地区中按抽中的次数平均分配，如抽中一次，该地区的样本量为404；如被抽中两次，该地区的样本量为808。

在地区少于5个的省中，各省的样本量（2020）在各地区中按地区的住宅电话数目的比例分配。

第二阶段：抽取电话局号

对于抽中的地区，通过有关部门获得了该地区的所有电话局号资料。

电话号码由区号、局号和后4位数字（有极少部分地区，交换机还没有更新，电话号码除了局号，后面是3位数字）构成。抽中的地区中有些地区是统一的区号，有些地区可能还有不同的区号。每个区号下有很多局号，局号相同，则地理位置可能比较相近，为使样本的代表性更强，每个地区的所有局号都被抽取。平均分配每个局号下的样本量。

第三阶段：抽取电话号码

由于各局号下的住宅电话号码未知，利用随机数字表，随机产生4位（或3位）随机数，与区号、局号一起构成号码。



很显然，这样获得的号码可能是空号（没有这个电话号码）、住宅电话和非住宅电话。

由于各个局号下所包含的电话号码数目和住宅电话号码数目都是不同的，所以不同局号下的住宅电话号码被抽中的机会不同，并且概率事前未知，为保证各省的样本近似于自加权样本，理论上此阶段要采取事后加权的方式，事后加权系数应由每个局号下住宅电话数目占全部有效电话号码数目的比例来确定，但实际上这个数据没有办法得到，只能利用拨打电话时获得的有关记录（如每个局号下有人接听的电话中住宅电话的比例）去估计加权系数，但这样做的误差可能很大，因为无法判断无人接听的电话是否是住宅电话。

在实施过程中，每个抽中地区的所有电话局号都被抽取了，由于每个地区下的局号数目都很大（一般都有数百个），也就是说，样本单元的分布遍布了所有的局号（相当于入户面访调查时抽取了该地区的所有居/家委会），因此，虽然各省的样本不是严格意义上的自加权样本，但是，将其按近似的自加权样本处理，误差可能反而会大大小于按近似估计的加权系数来计算所造成的误差。

第四阶段：抽取调查对象

电话调查方式在中国是新生事物，人们对此还不熟悉，为了便于访问实施，提高访问成功率，把接听电话的人作为被访对象，先询问其家庭基本状况和个人是否上网、个人背景资料和家庭其他成员的最简要资料。如果他（她）不上网，但家中有人上网，则再随机抽取一名上网的成员来接听电话，回答有关上网的问题以及他（她）的个人基本资料，以获取尽可能多的有用信息。

四、"大学生"子总体的抽样

略

五、总体主要目标量及其精度的估计

本调查的目标量可以分成3类：

1. 总量：比如说目标总体中网民的总数；

2. 比例：本调查的总体目标量大部分都是以比例的形式出现的，如各省中网民人数占目标总体的比例，不同性别、不同文化程度、不同年龄的人群的上网比例等；

3. 均值：比如说平均的上网时间、平均的自费上网费用等。

以下针对比例、总量和均值指标给出了估计公式和精度的估计式。略。

六、结束语

略



第八章　调查问卷的设计

记者在确定了研究题材，提出了研究问题或研究假设之后，就需要选择适当的研究方法。在各种社会科学研究方法中，经常为精确新闻报道记者所采用的一种方法就是使用问卷调查的研究方法。

第一节　问卷的作用与类型

一、问卷调查的功能

问卷调查是一种收集资料的方法，通常涉及抽样、问卷设计和实施访问等步骤，主要的目的是借助问卷这一工具，访问少数具有代表性的人或团体，对访问结果进行科学的推断，从而了解大众的生活状况、日常活动，以及他们对各种人物、事件等的意见和态度。从事精确新闻报道的记者，可以使用问卷调查的方法来收集下列4种信息：

1．人口统计特征

人口统计特征（Demographic Characteristics）指人群的基本背景资料，如性别、年龄、职业、收入、教育程度、婚姻状

况、居住地区、宗教信仰等。

通过问卷调查可以得到最新的人口统计资料，对于这一部分信息的收集通常是问卷调查中必备的一部分。通过对这些人口信息进行数据分析，可得到不同特征人群的情况。

2. 社会与经济状况

社会与经济状况包括社会治安、交通环境、健康及医疗设备、教育问题、经济发展、生活品质等。

3. 各类活动

各类活动所包含的信息范围非常广泛，如个人爱好、生活习惯、参与的主要休闲娱乐活动、旅行郊游计划、消费情况、收视阅读习惯，以及各种社会活动、政治活动等，大都可以采用问卷调查的方式来获取有关信息。

4. 意见与态度

人们对于各种问题、人物、事件的意见与态度，是新闻界最重视的一种信息。例如，大众对于各种政策制度的评价，对政府政绩的满意程度，对各种社会事件的反应等。这些都是新闻界所关心的问题，也是从事精确新闻报道的记者最常探讨的问题。

二、问卷的作用、类型与构成

以从被访者处获得信息为目的的一系列格式化的问题，称之为问卷。在数据收集过程中，问卷起着核心的作用。在问卷调查中，由于数据是依赖于问卷才得以收集的，因而问卷是影响数据质量的主要因素。设计一份完善的问卷不是一件容易的事情，问卷设计人员除了要具备传播学、新闻学、统计学、社会学、经济学、心理学、计算机软件应用等相关方面的知识外，还需要有一定的技巧和经验。



1．问卷的作用

在调查中，问卷具有以下几方面的作用或目的：

（1）将所需的信息翻译成被调查者可以回答，并且愿意回答的一系列具体的问答题。

我们知道，在社会科学研究中，对于概念的研究，是通过对概念的操作型定义的研究来实现的。而在问卷中，则是通过一个或一组问题来体现。例如，迈尔（Meyer，1973）在对黑人的调查中，对"好战倾向"所下的操作型定义，是通过对以下8个问题的回答得分来体现的：①

问题一：依您之见，政府推动黑白融合的速度是太快、太慢，还是刚好；

问题二：您希望看到更多的示威活动，还是较少的示威活动；

问题三：黑人努力工作也能和其他人一样容易出人头地；

问题四：黑人应多花一些时间祈祷，少花一些时间示威；

问题五：坦白说，我害怕参加争取民权的示威活动；

问题六：一家餐厅的老板如果不愿意服侍黑人，应该可以不服侍；

问题七：在黑人获得平等权利前，他们应显示自己值得接受平等权利；

问题八：房地产所有者如果不愿意把房地产卖给黑人，应该可以不卖。

在问卷中，有以上8个问题的态度指标，藉以研究"好战倾向"这一概念。这样，使得概念成为调查者能够回答的

① ［美］Meyer·Philip. Precision Journalism: A Reporter's Introduction to Social Science Methods. Indiana University Press. 1973。

问题。

同时，由于采用问卷方式，能将人们的态度、观点、行为、看法等定性认识转换为定量数据，因此很容易对得到的信息进行统计处理和定量分析。

但是要设计好被调查者可以并愿意回答的问题，而且这些问题还能生成所需信息，是件比较困难的事，两种表面上类似的提问方式可能会导致不同的信息。因此，这是一件具有挑战性的工作。

(2) 问卷必须促进、激发并鼓励被调查者参与、合作并完成调查。

没有全部完成的问卷，其作用是有限的。在设计问卷时，研究者要注意问题的措辞、逻辑顺序等，努力使被调查者的疲劳、厌烦减至最小，尽可能减少无回答情况的出现。

(3) 问卷应该使回答误差减至最小。问卷可能是形成回答误差的主要根源，使这种误差减至最小是问卷设计的一个重要目的。

2. 问卷的类型

问卷是一种调查工具，在调查中使用的问卷类型随所研究问题、对象和方式的不同而不同。

(1) 按照调查的方式分类

按照调查的方式分，有“自填式问卷”和“访问式问卷”两种。自填式问卷是通过面访或邮寄的方式，将问卷交到被调查者手中，由被调查者自行填写；访问式问卷是在面访或电话调查中由调查员将题目念给被调查者听，再由调查员把被访者的回答填写在问卷上。

(2) 按照结构分类

按照问卷的结构分，有“无结构型问卷”和“结构型问



卷”两种。无结构型问卷指对问卷中所提的问题，在组织结构上没有加以严格的设计和安排，只是围绕着研究目的来提一些问题。问卷中没有规定答案的选择范围，被调查者可以按照自己的意愿自由发挥。因此从形式上说，一般都是开放式的。例如：您对物价改革有什么看法？您对夏季作息时间如何认识？

结构型问卷是根据研究目的和主题精心设计的，有具体结构的问卷。问题按照一定的提问方式和顺序进行安排。例如，在广播电视的听众观众调查中，整个问卷按照受众的视听习惯、视听时间、视听兴趣、收获、观念态度等几个方面来安排问题。

(3) 按照回答问题的形式分类

按照回答问题的形式分，有“开放式问卷”和“封闭式问卷”两种。开放式问卷通常由开放式问题构成，提出问题时，不给被访者提供回答项，而由被访者使用自己的语言来回答问题。

封闭式问卷是将问题的内容和可供选择的答案做了精心的设计，被调查者可以在所规定的答案范围内进行选择。例如：

请问您平时去电影院看电影吗？

1. 经常看　　2. 偶尔看　　3. 从不看

封闭式问卷既便于被调查者回答，又便于资料的统计处理，是目前在调查研究中经常采用的方式，但在问卷的最后，有时会加上一个或几个开放式的问题，用于收集一些封闭式问题中未包含在内的更加生动、形象和具体的资料。

3. 问卷的构成

问卷一般可以看做由3大部分组成：封面信、正文部分和结尾部分。

问卷的封面信应包括的内容：提供调查的标题或主题；说

明调查的主办单位和自己的身份；说明调查的目的、内容和填写方法；说明完成问卷的重要性；征得被调查者的配合；确保被调查者提供有价值的信息；说明如何使用调查数据；说明保密的程度和数据的协议；在邮寄调查中，提供返回地址和返回日期。

封面信的语气应该是亲切、诚恳和礼貌的。问卷的封面信十分重要，因为在问卷调查中，调查的成功完全依赖于被调查者的自愿合作。大量调查的实践表明，几乎所有拒绝合作的人都是在开始接触的前几秒内就表示不愿参与的。被调查者一旦开始回答，就几乎都会继续并完成，除非在非常特殊的情况下才会终止。因此，一开始的介绍是绝对重要的。如果封面信的介绍和引入很恰当，回收率将会增加，调查的可靠性和有效性也会得到加强。否则拒答和不答的现象将会造成数据收集时间的拖长以及费用的增加，导致误差和偏差的产生。

问卷的正文部分实际上也分了3个部分，第一部分是开场问题，起着鼓励被调查者参与调查的作用。应向被调查者了解的最一般的问题，这些问题应该是适用于所有被调查者，并且能够很快、很容易地回答的问题。特别要避免那些可能会吓坏被调查者的问题。在这一部分不应有任何难答的或敏感的问题。

正文的第二部分是主要的问题部分，它包括涉及调查主题的实质和细节的大量问答题目。这一部分的结构组织安排要符合逻辑性并对被调查者来说是有意义的。从一个主题到另一个主题的转化要平稳自然，不要发生突变或大的跳跃，使得被访者难以接受。

正文的第三部分一般包括两种内容的问题：其一是敏感性的或复杂的问题，以及测量被调查者的态度或特性的问题；其



二是人口状况或基本情况，例如，涉及人口统计特征的问题，如被调查者的性别、年龄、教育程度等等。

问卷的结尾一般可以加上一个或两个开放题，给被调查者一个自由发表意见的机会。然后，对被调查者的合作表示感谢。

第二节　问卷设计的基本步骤

一、问卷调查面临的困难

一份成功的问卷应该达到两个目标：一是设法取得被访者的合作；二是能将所要调查的问题明确地传达给被访者，以取得真实、准确的回答。但在实际的调查实施过程中，由于被访者的个性不同，他们的教育水准、理解能力、道德标准、宗教信仰、生活习惯、职业和家庭背景等往往存在巨大差异；再加上调查员本身的专业知识与技能高低不同，这些都会对调查的结果产生影响。在问卷调查中，可能会遇到这样的困难：

第一，被调查者不了解或是误解问题的含义，造成无法回答或答非所问。

第二，被访者虽然能够了解问题的含义，愿意回答问题，但记忆不清应有的答案。

第三，被访者了解问题的含义，也具备回答的条件，但不愿意回答，即拒绝回答。导致拒答情况出现的原因很多，当被访者对问题毫无兴趣，或被访者对问卷有畏难情绪，对提问内容有所顾虑时，都可能会不愿意回答。

第四，被访者愿意回答，但没有能力回答，比如，被访者不善于表达意见，不适合回答，或不知道该如何回答等。

为了克服上述困难，达到问卷的两个主要目标，问卷设计应遵循目的性原则、可接受性原则、简明性原则和匹配性原则，并按照一定的程序来设计问卷。

二、问卷设计的基本步骤

问卷设计的基本步骤见图 8-1 所示。在实践中这几个步骤是相互联系的，在设计问卷的过程中，有些步骤可能需要重复或循环。

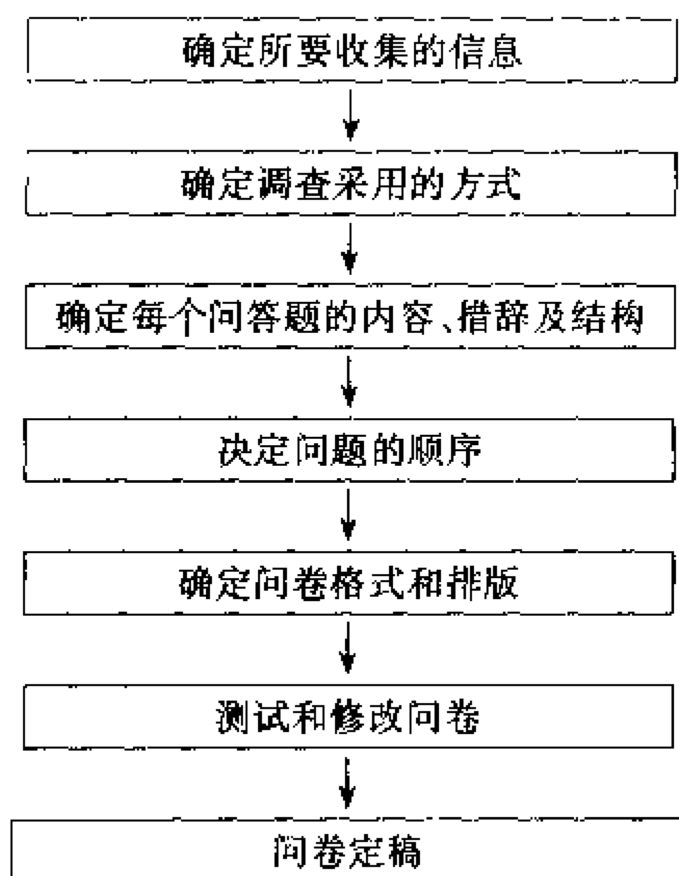


图 8-1

1. *确定所要收集的信息*

问卷设计的第一步就是确定所需要收集的信息，在这个过程中，需要考虑研究问题和研究假设，并将研究问题具体化、



条理化、操作化，变成一系列可以测量的变量或指标。

例：某报希望在北京市民中进行一项关于“电子宠物有多热门?”的民意调查，该报希望回答以下研究问题：

(1) 北京城区居民知道“电子宠物”的比例；

(2) 在知道“电子宠物”的居民中，赞成与反对“饲养电子宠物”的比例；

(3) 在知道“电子宠物”的居民中，“饲养电子宠物”的比例；

(4) 性别、年龄、教育程度是否与以上问题相关。

要回答以上研究问题，问卷中的问题至少要涵盖以下方面：

(1) 北京城区居民是否知道“电子宠物”；

(2) 知道“电子宠物”的居民，是否赞成“饲养电子宠物”；

(3) 知道“电子宠物”的居民，是否“饲养电子宠物”；

(4) 被访者的年龄；

(5) 被访者的性别；

(6) 被访者的教育程度。

以上方面在问卷中遗漏任何一个，都将导致无法完整上述研究问题。所以研究者在列出问卷问题所需涵盖的重点，准备拟定研究问题时，最好先将所有可能和上述几个方面有关的问题都写下来，然后再把没有关系，或是比较不重要的问题删掉。

当然，拟定问卷上的问题，最好能参考过去的类似研究以及研究问卷上的问题。比如，国外新闻界进行选举预测时，就可能会考虑采用或修正盖洛普选举调查的问题。

另外，对研究总体也要有一个清楚的概念，了解被调查者

的社会阶层、行为规范、社会环境等社会特征；文化程度、知识水平、理解能力等文化特征；动机、行为等心理特征，并以此作为拟定问卷的基础之一。

2．确定调查采用的方式

不同类型的调查方式对问卷设计是有影响的。在面访调查中，问卷由访问员进行管理，被访者可以与访问员面对面地交谈，因此，可以询问比较长的、复杂的和各种类型的问题。在使用电话访问的方式时，只限于问一些短的和比较简单的问题。邮寄问卷是被调查者自我管理的问卷，因此问题也要简单一些，并给出详细的指导语。所以在设计问卷之前，要先确定调查采用的方式，并依此方式进行问卷设计。由于电话调查在精确新闻报道中扮演着越来越重要的角色，所以对于电话调查问卷的设计在后面的内容中，还将详细加以介绍。

3．确定每个问答题的内容、措辞及结构

（1）确定每个问答题的内容

问卷设计要紧紧围绕研究假设进行，问卷中必须问什么，不必问什么都将受到研究假设的制约。在开始的时候，不妨将与研究假设和研究问题有关的问题都写下来，然后问自己：这个问题有必要吗？问卷中的每一个问题都应对所需要的信息有所贡献，如果从一个问题得不到可以满意使用的数据，那么这个问题就应该取消。问题的个数应该限制在最少数量，每个问题都必须有其纳入问卷的原因。

在确定某个问题是必要的以后，还要避免“一题两问”的情况出现，例如：“您支持政府关于环境和教育问题的立场吗？”被调查者可能很支持政府对于环境问题的立场，但不支持其对教育问题的立场，那么被调查者将无从回答这个问题。在这种情况下，应该用两个不同的问题进行询问，即“您支持



政府关于环境问题的立场吗”和“您支持政府关于教育问题的立场吗”。

(2) 确定每个问题的结构

调查问卷中的问题从结构上讲，分为开放式问题、封闭式问题和混合式问题。

开放式问题不给被访者提供回答项。被访者使用自己的语言或者通过提供精确的数字答案来回答问题。

例如：您为什么不赞成在北京实行夏时制？

作为北京市民，目前您最关注的问题是什么？

请问您的年龄有多大？

开放式问题的优点是被访者可以充分自由地按照自己的想法和方式回答问题，不受限制，有利于发挥被访者的主动性和想像力。因此开放题收集到的资料往往比较生动、具体，而且信息量大。开放式问题可以应用于探索性研究中，或用在结构式问卷的结尾等处，如“您对我们刚才讨论的主题还有什么补充意见吗？”这就确保被访者有机会将他们认为与问卷中所列问题有关的每件事物都包括在内。另外，开放式问题还可以用来对封闭式问题进行补充。

开放式问题的缺点可从两方面进行描述，对被访者来说，开放式问题的压力更大，因为被访者必须在没有回答项帮助的情况下确定问题的内容。在自填式问卷情况下，由于被访者需要构思答案并记录下来，他们需要更多的时间来完成问卷。如果问卷是访问员控制式的，对被访者和访问员来说，都仍然要花费较多的时间，因为必须将所有回答记录下来。由于时间关系和缺乏心理准备，被访者往往放弃回答，使开放题的回答率不高。

对研究者来说，一般要求访问员逐字记录答案。这通常是

很困难的，开放式问题会导致调查过程的费用增高、时间加长、容易出现误差。在收集开放式问题的答案之后，通常要对它进行数字编码。这一过程的费用也比较高，又很费时间，并会使最终结果产生编码误差。因此开放题难于进行定量整理和分析。

封闭式问题规定了一组可供选择的答案和固定的回答格式，被调查者只要或只能从中选定一个或几个现成的答案。封闭式问题量表的形式可以是多种多样的。

例如：请问您平时去电影院看电影吗？

1．经常看　2．偶尔看　3．很少看　4．从不看

请问您对您目前的生活状况是否满意？

1．很满意　2．比较满意　3．无意见　4．不太满意　5．很不满意。

封闭式问题的优点在于被调查者只需简单地选择合适的回答项，这样他们就能够更快、更容易地回答问题。对于研究者来说，封闭式问题由于答案标准化，易于进行各种统计处理和分析。封闭式问题的缺点是设计比较困难，要花费很多精力，确保问题答案项之间是排他的、穷尽的，一旦设计有缺陷，被调查者可能就无法正确回答问题，从而影响调查质量。另外，封闭式问题的答案项可能使所研究的问题简单化，而被调查者只能在规定的范围内回答，而无法反映他（她）的真实的、复杂的想法。

混合式问题是指问卷上所问的问题，一方面事先设计答案项，使被调查者可以从中选择，另一方面，也给被调查者表达个人意见的机会。比较典型的混合式问题是在封闭式问题的答案项中，加一项“其他”，并要求被调查者详细说明。

例如：您家平均每月支出中，花费最多的是：



1．食品　2．服装　3．娱乐　4．日用品　5．交际　6．其他（请注明）＿＿＿＿＿＿

在问卷中采用混合式问题，一方面可以保证封闭式答案项的穷尽，另一方面是让被调查者在回答封闭式问题时，仍然有机会自由回答。

精确新闻报道采取混合式问题，至少具有下列两个优点：

首先，记者可以把从混合式问题中所得到的多数被调查者的意见，配合数字比例说明，呈现在新闻写作中，使新闻不致完全以数字为主，因而增加新闻的可读性。

其次，如果被调查者答出特别有趣或突出的言论，记者在进行精确新闻报道写作时，可以用直接引句加以引用，使精确新闻具备轻松、生动的一面。

（3）确定每个问题的措辞

问卷设计者要尽量使问题对被访者来说既清楚，又包含设计者的本意。如果被调查者能清楚地理解所用措辞的含义，调查数据将是高质量的。如果被访者清楚地理解了所问的内容，他们也将更愿意并能够提供信息。此外，被访者按照调查设计者所希望的方式理解调查问题，这是很基本的要求。

问题的措辞在很多方面都会影响被调查者的回答，即使对措辞做微小的改动，也可能造成调查结果上的显著差别。下面是一个这样的例子[①]，该研究在德国进行，研究中使用两个不同版本的问题对全职家庭主妇进行访问，每个版本分别用于一半的样本。第一个版本的问题是："如果可能的话，你希望有一份工作吗？"第二个版本是："你是更喜欢有一份工作，还是

① 资料来源：Norman M. Bradburn Seymour Sudman《Polls &Surveys: Understanding what they tell us》Jossey-Bass Publishers 1988。

是喜欢只做家务呢?”研究结果表明，在使用第一个版本的样本中，19%的被调查者回答她们不愿意去工作，而使用第二个版本的样本中，表示不愿意去工作的比例达到68%。

所以，问题的措辞非常重要，要使被调查者和研究者对问题的意义理解完全相同，否则结果会产生严重的偏差。

如果出现下列情况，一个问题的措辞就有可能歪曲调查结果，提供模糊的数据：

- 被访者不理解问题中措辞的含义；
- 对措辞的理解与原意不同；
- 不熟悉问题中措词所表达的概念。

为避免这些问题，必须记住一些基本准则：

①措辞要简单、明确、易懂

与被调查者进行沟通的最佳途径是使用简单、日常的措辞，并确实使所有的术语都适合被调查总体。在设计问题时，要经常考虑被访者的语言技能。尽量选择每个人都很容易理解的措辞。

例如：你以前接种过肺炎球菌疫苗吗?

普通人群中大多数人不知道医学术语——肺炎球菌，因而回答问题比较困难。结果，一些人可能无法提供准确的答案。

如果一定要在问题中使用新的或复杂的概念，必须在问卷中进行定义，以使所有的被访者对问题的理解是相同的。

另外问卷中使用的词汇对被调查者来说应当具有惟一的意义。有些似乎是明确的词汇，实际上对不同的人是有不同意义的。

例如：请问您经常看电影吗?

每个被调查者对“经常”这个词的理解可能是不同的，是一周一次，还是一月一次等。



②确保所提的问题对被调查者是适用的

例如：请问您一天抽多少支烟?

对于这样的问题，如果被调查者根本不吸烟，那么这道题他（她）将无法回答。解决的办法是在这道题目之前加一道过滤性问题，即：

请问您抽烟吗?

如果被调查者回答“是”，则继续提问刚才的题目。

③问题要具体，界定明确

调查问题的措辞要尽量具体，以确保被调查者能确切理解问题的意思。一定要清楚地说明：

- 问题适用于谁；
- 问题界定的基准期是什么；
- 回答中应包括或不包括的信息；
- 应在答案中提供单位。

乍一看，下面这个问题可能很简单、很直接：

请问您的收入是多少?

再仔细一看，这个问题实际上并不那么容易回答。首先，您指的“谁”不明确。研究者是想了解被调查者的个人收入，还是他或她的家庭收入这要搞清楚。其次，要求被调查者提供收入信息的基准期是什么? 上周、上个月、还是去年? 最后，被调查者应计算什么收入? 只是工资? 还是包括奖金? 以及其他来源的收入? 或者别的什么?

下面提问的问题是两种较好方法：

请问去年，在交税和各种扣除之前，您家各种来源的收入总共是多少元?

或：

请问去年，在各种扣除之前，您全家的总收入是多少元?

（包括来自工资、薪金和所有其他来源的收入）

④避免引导性或带有倾向性的问题

引导性问题是指暗示或引导被调查者选择某一特定答案的问题。换言之，问题措辞方法会影响对问题的回答。引导性问题能误导调查回答并影响调查结果。

例如：引导性问题

请告诉我，你对“旅游业有益于本县，应该促进”这种说法是同意、不同意、还是没有意见？

这样的问题暗示了旅游业有益于本县，引导被调查者回答“同意”，因此使研究者无法了解被调查者的真实想法。比较好的提问方式应该采用下面的中性问题：

请告诉我，你对“应该促进本县的旅游业”。这种说法是同意、不同意、还是没有意见？

⑤避免在问题中使用双重否定

在问题的句子结构中，应避免双重否定。请考虑下面的例子：

您是赞成、还是不赞成学校不允许学生经商的规定？

要回答这个问题，被调查者就必须搞清赞成不允许学生经商，也就是反对这样做；同样，反对不允许学生经商，也就是赞成这样做，这要求动一定的脑筋。

这个问题很难回答，因为它包含一个双重否定：反对和不两个否定词。包含双重否定的问题会让被调查者感到迷惑，不知道如何回答，可能做出与真实想法相悖的回答。最好的办法就是重新改述问题，让它只包含一个否定词。

⑥考虑问题的时间性

时间过久的问题容易让人遗忘，而且会加重被调查者的回答负担。比如下面的问题：



过去30天内您花了多少小时看电视？

一般情况下，被调查者很难回答这个问题，即使回答了，答案的可信度也不会高，从这样问题得到的数据对研究者的意义也不大。一般来说，人们对最近发生的事情会记得清楚一些。研究者可以采取这样的提问方式：

请问上周您平均每天看了多少小时的电视？

缩小时间范围可以使被调查者回忆起来比较容易，回答也会比较准确。

⑦设法减轻敏感问题或隐私问题的影响

被调查者对于敏感、隐私或有威胁性的问题，可能只选择那些最能满足自己自尊心的，或符合社会规范的答案，而不表达他们的真实想法或真实情况。各地风俗和民族习惯中忌讳的问题、设及个人利害关系的问题、个人隐私问题等都属于会令被调查者难堪、禁忌或敏感的问题。

例如：你有没有在酒后驾驶过机动车？

你的体重是多少？

你认为自己是超重、太瘦，还是正好？

您是否有逃税行为？有过几次？

对于这类问题，被访者往往出于本能的自卫心理，容易产生种种顾虑，不愿意回答或不予真实回答，而且还会引起被调查者的反感，由此影响对其他问题的回答。但是对于精确新闻的研究主题来讲，这样的问题有时候是不可避免的，因此在设计这类问题时，要尽量注意提问的方式、方法和措辞。仔细地措辞能够帮助减轻敏感问题对调查结果的影响。在实际操作中，可以考虑采用以下几种方式：

1）“每个人”法

用问题的措辞来说明这种敏感的行为并非不常见，从而有

效地引入敏感性话题。比如，可以使用“许多人”或“大多数人”这样的词汇来引入。

2）释疑法

即在问题之前写一段功能性文字，用来消除被访者的顾虑。

3）系列问题或热身法

有时候，可以先用几个问题来减轻被调查者对敏感话题的敏感程度，而真正的问题在热身问题之后出现。或者可以将敏感问题“隐蔽”在一组被调查者容易而且愿意回答的问题之中。

4）转移法

即把本应由被调查者根据自己的实际情况填答的问题，转移到由被调查者根据他人的情况来阐述自己的想法。

例如：想像有两个人在讨论各自的生活方式。其中一个人认为，一个人应该履行自己的义务；另一个人则说，个人愿望的满足才是最重要的事情。他们两个都问您的观点是什么？请问……①

5）回答范围法

对于一些个人的特殊信息，如年龄、收入或不良行为频率等方面的敏感信息，如果让被调查者回答一个范围而不是回答确切数字的话，他们可能更愿意回答。

例如：去年，您个人在纳税前的全部收入是多少？（包括工资、奖金和所有其他来源的收入）

a．低于500元

---

① 资料来源：北京广播学院调查统计研究所1999年《现代科技与生活调查》项目。



b．501～1000元

c．1001～2000元

d．2001～3000元

e．3000元以上

4．决定问题的顺序

问题的顺序应该按照鼓励被调查者完成问卷，并保持他们的兴趣来设计。这种顺序还应该能够便于被调查者回忆，并使被调查者觉得有道理，向被调查者反映出顺序的逻辑性，同时集中调查主题，使得调查主题很顺畅地从一个到另一个，相同主题的问题放在同一部分。

在设计问题的顺序时，应考虑遵循以下几个原则：

（1）先简单问题，后复杂问题

（2）先基本信息、分类信息，后身份信息

一般我们从问卷调查中获取的信息可以分为3类：基本信息、分类信息和身份信息。其中基本信息直接与研究问题相联系；而分类信息主要包括社会经济和人口状况方面的内容，用于对调查对象分类；身份信息指姓名、地址、电话号码等。当然多数民意调查是不记名的，但在必要的情况下，也可以要求被调查者自愿地提供他们的身份信息。一般来说，获取信息的顺序是基本信息在前，接着是分类信息，最后才是身份信息。

（3）先一般问题，后敏感问题

将敏感性的问题尽量放在问卷的后面，因为经过一段时间以后，被调查者最初的不信任感已经克服，友善感已然建立，问题的合理性也很明显，这时被调查者可能更愿意提供信息。

（4）先封闭式问题，后开放式问题

一般来说，开放式问题需要被调查者有思考的时间，回答的难度也比封闭式问题要大，如果将开放式问题放在前面，可

能会使被调查者产生畏难情绪。

(5) 先综合性问题，后具体问题

处在问卷前面的问题可能会影响后面的问题。一般来说，综合问题要放在具体问题之前，这样可以防止由具体问题造成综合问题答案的偏差。

例如：在您选择旅游的方式时，哪些因素是重要的？

在您选择旅游的方式时，费用这个因素的重要性如何？

显然第一个问题是一个综合性的问题，而第二个问题是具体的，如果我们将这两个问题调换顺序，那么被调查者就会受到“费用”这一因素的提示，在回答综合性问题时更可能考虑“费用”这个因素。

(6) 注意问卷的逻辑顺序

问题在问卷中的顺序要体现逻辑性，有关相同主题的问题放在同一部分，可以加上标题，如环境关注、健康问题等，来说明这一部分的主题。从一个主题变换到另一个主题时，可以加一些简短的说明，如“现在让我们来谈谈环境问题，我指的是您周围的自然环境，空气、水等等”，以使过渡自然。

对于分叉问题要特别仔细，分叉问题常常伴随着跳答模式，为了不产生错误，这里介绍一种框图设计法。在问卷设计之初，就先画出整个问卷的各个部分及前后顺序的框图。然后再具体写出每一部分中的具体问题，并安排好这些题目间的相互顺序，在相应需要跳答的题目处给出明确的指导。

5. 确定问卷的格式和排版

问卷的格式、问答题的位置、题目之间的距离等等因素都可能对调查的结果产生影响，有关的实验结果表明，放在每页顶部位置的问题一般比放在底部位置的问题更容易受到关注。



前面我们提到在每个主题部分可以加一个标题，以使问卷条理更清楚。在对设计完成的问卷排版时，还应考虑很多因素，包括字型、字体等。在问卷中，要保持问题、指导语、标题和过渡性说明的一致性和连贯性，这可以通过使用不同的字型和字体来区别这些发挥不同作用的文字来实现，以便使被调查者或访问员很容易地辨认这些文字。例如，可以使标题的字体比题目或答案项的字体大一些等；有时也可以使用图表增强问卷的活泼性。

问卷排版时，最好每个问题都在单页纸上，不要将一个问题（包括可选择的答案项）分开在两页纸上。

6. 测试和修改问卷

在问卷设计完成之后，要对问卷进行测试。由于被调查者对问题作出反应是一个复杂的过程，被调查者必须首先理解问题，然后进行思考或搜寻回忆，以便找出所需要的信息。在找到这些信息之后，还要考虑什么是正确的答案，以及自己愿意在多大的程度上披露这一答案。这些步骤中的每一个环节，都可能是回答误差的来源。测试问卷，旨在确定问题中存在的困难和错误，确定问题的顺序是否影响了对问题的解释，确定指南是否清楚，以及被调查者或访问员对问卷的外观感觉如何等。

对问卷初稿测试的方法比较常用的有两种：

一种是在一个小样本中对问卷进行测试，用于测试的样本也应该取自实际调查的总体，样本量并没有一致的规定，一般在30个左右即可，测试问卷最好采用面访的形式，即使是实际的调查将采用电话或邮寄的方式，因为这样可以观察到被调查者的反应和态度，问卷设计者和主要的研究人员应该参与一些测试面访。测试所用的调查员应该包括有经验的和新访问

员，以便确认访问能力不同的访问员，在对问卷进行操作时，可能出现的问题。

问卷初稿测试的另一种方法是主观评价的方法，即将设计好的问卷初稿，请该研究领域的专家、研究人员以及典型的被调查者进行评价，找出不妥之处。

只有经过测试，找出问题并进行了修改，才能把问卷用于正式调查。

7. 问卷定稿

问卷定稿以后，如果不是采用 CATI（计算机辅助电话调查）或 CAPI（计算机辅助面访调查）的访问方式，那么就要将问卷印刷。问卷如何印刷和装订也可能影响到调查结果，问卷应该用质量比较好的纸张印刷，并且正规地装订，给人一种“专业性”或“职业性”的感觉。

三、问卷设计中的其他问题

1. 减少测量误差的技术

这里所说的测量误差是指对问题的真实回答与被访者对问题的回答之间的差异，由于问卷是收集数据的手段，因而在设计问卷时使测量误差最小化是非常重要的。我们可以识别测量误差的来源，同时应用有关技术来减少这类误差的影响。

比如，我们前面提到的几种处理敏感问题的技术，就可以在一定程度上减少被调查者的回答误差；在被调查者回忆过去的行为或时间时，也可能产生误差，在这种情况下，可以采用缩短基准期的方法来解决；问题的长度对测量误差的产生也有一定的影响，有研究成果表明，较长的问题可能使得被调查者有更多的时间来思考，并能够提供更完整的答案。

2. 变量设计



问卷设计应该紧紧围绕调查研究的目的、主题以及研究假设，为了配合整个研究，在问卷设计时要明确有哪些自变量、中间变量和因变量。围绕每个变量设计一个或多个问题。考虑到将来可能要分析研究的问题，在设计问卷时都要配以适当的变量，尽量不要遗漏。

3. 问卷不能太长，所问项目不要太多，可要可不要的问题一定要删掉。

4. 问卷中所用语言要口语化，符合人们日常提问和交谈的习惯，避免过于书面化。

5. 有的时候，问题回答项的排列顺序会对被调查者的回答产生影响，因此可能需要有不同排列顺序的几个版本的问卷来消除这种影响，或者在面访或电话访问中，通过控制访问员轮换回答项的顺序，来消除这种顺序偏差。

第三节　问卷中常用题型和量表的设计

在调查问卷中，每个问题都对应于一个或多个变量，测量这些变量的值，需要一定的工具，这些工具就是所谓的量表。通俗地说，量表即是我们在问卷中呈现的封闭式问题的题目与答案。

由于我们的问卷中大多数问题是封闭式问题，因此题型与答案的设计必须经过多方面的细致考虑。

一、量表的类型

前面我们提到变量的测量级别主要有4种：定类变量、定序变量、定距变量和定比率变量。与此相对应的量表也有4类，分别是：定类量表、定序量表、定距量表和定比率量表。

1．定类量表

定类量表又称为类别量表，或名义量表，定类量表给出的数字仅仅可以用来识别被调查对象的性质或特征，由此实现对被调查对象的分类，而并不反映被调查对象所具有某种特征的数量。例如，被调查者的性别可以用1表示“男性”，2表示“女性”。

对于用定类量表所测量的定类变量，可供选择的统计方法有计算频数，以及与频数有关的统计量和统计方法，例如，计算百分数、众数，做卡方检验等。

2．定序量表

定序量表又称为顺序量表，这种量表对被调查对象给出的数字表示个体具有某种特征的相对程度，指示被调查个体之间的相对位置，但不能指示个体之间具体的差值是多少。在民意调查中常用的定序量表有满意度排序、喜好程度排序等。

例如：请问您对《泰坦尼克号》这部电影的喜欢程度如何？

1．非常喜欢　2．比较喜欢　3．无所谓　4．不太喜欢　5．非常不喜欢

对于由定序量表测量的定序变量，主要可以应用以下几种统计方法：

（1）与定类变量类似的频数分析；

（2）与频数分析相关的统计方法；

（3）秩相关分析。

3．定距量表

定距量表又称为等距量表，定距量表不仅可以表示被调查个体之间的顺序关系，而且可以测量被调查个体之间的距离。

例如：请按照您的满意程度对北京市交警在以下方面的工



作打分（满分为 100 分）。

1．态度　2．……

在定距量表中，80 分与 40 分之差与 90 分与 50 分之差是相同的，但是不能说 80 分是 40 分的两倍，这是因为定距量表没有真正的零点。

对于用定距量表测量的定距变量，除了可以适用于定类变量和定序变量的全部方法外，还可以计算算术平均值、标准差、积距相关系数，其他常用的统计方法，如，方差分析、回归分析、因子分析等都是适用的。

4．定比率量表

定比率量表又称为等比量表，具有定类、定序和定距量表的所有性质，除此之外，定比率量表还具有绝对的零点。年龄、收入等变量都是可以用定比率量表来测量的。

例如：请问去年，在各种扣除之前，您全家的总收入是多少元？（包括来自工资、薪金和所有其他来源的收入。）

由定比率量表测量的定比变量，不但可以用来识别被调查个体或对个体分类、排序、比较差距；而且可以计算比值。比如，家庭全年收入 6 万元是全年收入 3 万元的两倍。

所有的统计方法都适用于定比数据。

二、问卷中的常见题型

1．二项选择题

二项选择题是封闭式问题中最简单的一种，用于将被调查者分成截然不同的两组。主要适用于对态度和意见的测量，互相排斥的两项择一问题，或简单的事实性问题。在这种题型的设计中，答案项只有两个，因此带有强迫性质。

例如：您家中有电脑吗？

1．有　　　　2．没有

对于“能够在有生之年跨入一个新的千年是十分幸运的事情”这种说法，您是否同意？

1．同意　　　　2．不同意

在第二个例子中，我们看到这样的问题和答案将被调查者简单地分为两类，使中立者意见偏向一方，不明确的态度明确化了。但是这种题型不能测量被调查者意见的程度。

2．多项选择题

多项选择题是调查问卷中最常使用的题型。这种题型是所问问题事先设计好两个以上的答案项，被调查者可以根据要求和自己的情况选择其中的一项或几项。由于事先设计的答案项不一定能穷尽所有可能的回答，所以在答案项的最后通常会设一个“其他”项。多项选择问题通常有以下两种情况：

(1) 多个答案的情况

例如：下列哪些类的电视节目是您所喜爱的？（选项不限）

1．新闻节目　　　2．体育节目　　　3．广告

4．综艺性娱乐节目　5．电影、电视剧　6．……

在上面这个例子中，被调查者可以从答案项中选择一个或多个答案。这种情况的多项选择题，其指导语除了“选项不限”之外，还有“限选3项”等等。在可以选择多个答案的情况下，每个可能的答案都对应一个待分析的变量。

(2) 单个答案的情况

当所设计问题的选择准则只定义了一个结果，要求被调查者从可供选择的答案中只能选择一项。例如，我们将上面的例子换一种问法：

请问您最喜欢看以下哪一类的电视节目？（只选一项）

1．新闻节目　　　2．体育节目　　　3．广告



4．综艺性娱乐节目　5．电影、电视剧　　6．……

这种问法定义了被调查者只能选择一项，即自己最喜欢的电视节目。

使用多项选择问题要遵循以下几点：

①全部的回答应当是能够分成离散的有限种类别，但类别不宜过多；

②类别的名称应当能够清楚地定义出一组离散的可供选择的答案；

③给出的类别应当是互斥的；

④给出的全部类别应当能够包含全部可能答案的 90% 以上，一般最后加一个“其他”类别，来穷尽所有可能答案。

3．顺位题（排序题）

顺位题最主要的特点是使被调查者选择的答案呈现出顺序与程度。在顺位题中，被调查者按照题目指导语的要求将全部或部分答案项进行排序。

例如：您认为决定个人收入高低的因素主要是什么？（请将选项的序号填在下面的空格上）

1．工龄　2．学历　3．岗位职务　4．工作态度

5．业务技术能力　6．与领导关系　7．成就贡献

8．风险责任

9．体力支出

第一位因素＿＿＿＿＿＿＿＿＿＿

第二位因素＿＿＿＿＿＿＿＿＿＿

第三位因素＿＿＿＿＿＿＿＿＿＿

在设计顺位题时，要注意选项不宜过多，过多容易分散，很难顺位。另外选项的排列位置也可能对被调查者产生某种暗示的影响。

三、问卷设计中的常见量表

量表作为一种测量技术，可以分为比较量表和非比较量表，其中比较量表包括配对比较量表、排序量表、固定和量表等；非比较量表有两大类，即连续评价量表和分类量表，精确新闻报道记者比较常用的是分类量表，包括李克量表、语意差别量表、斯塔普量表等。

1．配对比较量表

当需要在多种可能的选择之间进行简单的两两比较时，就可以采用配对的比较量表。

例如：下面的每对电视连续剧，都是近期电视台播出的，如果您不得不从中选择一个的话，请选择您比较喜欢的那一个。

A：1．红色康乃馨　　2．让爱做主

B：1．红色康乃馨　　2．像雾像雨又像风

C：1．让爱做主　　2．像雾像雨又像风

为了避免问题的回答过于困难，要比较的目标的数目不能太多。配对比较量表的主要缺点是当比较几对目标时，其结果可能缺乏“传递性”。从逻辑上说，如果一个人喜欢 A 甚于 B，喜欢 B 甚于 C，那么当 A 和 C 比较时，显然他（她）应该更喜欢 A。但利用配对比较量表进行测量时，研究者常常会发现被调查者提供的信息缺乏传递性。

2．排序量表

排序量表产生顺序的数值，要求所有项目中的每一项都相对于其他项目被排上了顺序。

例如：请您按照您心目中的重要程度对决定个人收入高低的因素进行排序。最重要的因素排为“1”，第二重要的因素排



为“2”，依次类推，直到将所有因素都排了顺序。（请将序号填在下面的空格上）

1. 工龄____________

2. 学历____________

3. 岗位职务____________

4. 工作态度____________

5. 业务技术能力____________

6. ……

排序量表的主要局限是无法测到绝对的标准和项目间的距离。对这种类型的量表来说，可以排序的项目的个数应该限制在10个之内，否则会造成被调查者的回答困难。要求被调查者进行排序时，应当清楚地陈述一个判断准则，使得所有的项目都能以相同的标准得到排列。

3. *李克量表*（Likert）

李克量表是在民意调查中经常使用的一种量表。本书第二章第四节对李克量表作了介绍，并给出了李克量表的例子。

研究人员采用李克量表来测量人们的态度或观点是十分普遍的，因为这种量表功效高，形式简洁，同时可以计算各题项的总分，来表示对某个主要问题的态度的总指数。在设计李克量表时要注意，李克量表应该是围绕问题的“典型”观点的一组说法，而不是一两个说法，各种说法应该是分散的，能够代表关于所研究问题的一个比较宽的范围。尤其注意在设计量表时，要让一半的项目向着肯定的方向，另一半问题向着否定的方向，以消除回答偏差。

4. *语意差别量表*（Semantic Differential）

语意差别量表是研究者用一对互为反义词的形容词，来测试某个实体在被调查者心目中的形象。

使用语意差别量表，研究者必须首先选择一系列从不同维度描述所研究对象的形容词，以及确定每个形容词的反义词，之后就可以构造语意差别量表了。将形容词列于两端，大约一半的形容词将肯定的放在左边，另一半将否定的放在右边，中间可分为若干等级。然后要求被调查者就个人对所研究的概念、事件或物体的感觉，在量表上选出最接近自己感觉的位置。本书第二章第四节中给出了语意差别量表的设计实例。

语意差别量表的另一个用途是可以测量被调查者心目中理想的形象或属性水平，通过比较被调查对象心目中的理想形象可以确定不同的偏好模式。也可以将实际的与理想的形象加以比较。

设计语意差别量表的关键是研究者所选择的形容词应该能够规定那些对被调查者来说是重要的属性。如果忽略了某些重要的属性，就得不到被评价对象的完整轮廓。

5. 斯塔普量表（Stapel）

有时候如果想知道被调查者对某个问题是如何看待的，适合用什么样的形容词或词组来描述。斯塔普量表是在问题的指导语和一组形容词名单之间加一个数字量表，要求被调查者对每个项目都给出一个评分。例如，我们想要知道被调查者对自己生活的评价，就可以使用这种量表：

对于下面的每一个词，请从量表中选择一个适当的数字来描述您的生活，并将该数字写在对应词的前面的横线上。

完全不是　1　2　3　4　5　6　7　完全是

________乏味的

________充满挑战的

________有情趣的

________安定的



________愉快的

……

斯塔普量表也可以用来描述形象，但不像语意差别量表那样需要设计形容词及其反义词。

四、量表的设计和评价

在以精确新闻报道为目的的研究中，经常需要测量人们的意见、态度、看法或观念等比较抽象的东西，为此，必须将其中抽象的概念变的可操作化，设计一系列同质性的问题或指标，形成量表，加以测量。

1. 量表的设计程序

(1) 根据所要测量的内容或变量，收集大量与这一内容有关的问题，然后筛选出一组问题（一般为 10~30 个）作为初步量表。

(2) 确定问题的类别和记分标准。回答的类别可以分为 5 个等级，也可分为 7 个等级或其他测量等级。一般来说，在一个量表中，正向提问和负向提问应各占一半，以消除回答偏差。如果是 5 级量表，可用 1~5 记分。

(3) 试调查。从研究总体中找一个小样本来测试量表。

(4) 计算量表中各个项目的分辨力。所谓分辨力是指一个题目是否能区分出人们的不同态度或态度的不同程度。对于分辨力不高的项目要予以删除。

以李克量表为例，将试调查中得分最高的 25% 的人与得分最低的 25% 的人进行比较，计算出每个项目的分辨力。将量表中，每个人在每个项目上的得分加起来，计算每个人的总得分，按照总得分从高到低进行排序。

某项目的分辨力系数 = 得分最高的 25% 的人在该项目上

的平均分－得分最低的25%的人在该项目上的平均分

例如，某个由12个项目组成的李克量表，20个样本的回答情况按总分由高到低排列，见表8-1：

表8-1　　李克量表测试结果

被调查者＼项目		1	2	3	4	5	6	7	8	9	10	11	12	个人总分
总分高的25%人	样本1	5	4	5	4	3	5	4	4	3	5	2	5	49
	2	4	5	4	5	5	4	3	2	5	4	1	4	46
	3	5	4	3	3	4	5	4	3	4	4	2	4	45
	4	4	4	4	4	5	3	3	3	4	5	1	5	45
	5	5	5	3	2	4	4	3	4	5	2	2	4	43
	6	4	3	2	5	4	5	4	4	2	3	1	5	42
	7	4	4	4	4	2	3	3	4	4	3	2	4	41
	8	3	3	4	4	2	3	5	4	2	3	2	5	40
	…													
	14													
	15	2	4	2	3	2	2	3	4	4	3	1	4	34
总分低的25%人	16	2	2	4	2	3	3	2	1	4	2	2	5	32
	17	2	2	2	3	4	2	4	1	3	3	2	4	32
	18	1	3	2	4	1	3	3	2	1	2	2	5	29
	19	1	1	2	2	2	3	2	3	4	1	1	4	26
	20	1	1	1	2	1	2	1	2	3	2	2	3	21
高分组平均分		4.6	4.4	3.8	3.6	4.2	4.2	3.4	3.2	4.2	4.0	1.6	4.4	
低分组平均分		1.4	1.8	2.2	2.6	2.2	2.6	2.4	1.8	3.0	2.0	1.8	4.2	
分辨力系数		3.2	2.6	1.6	1.0	2.0	1.6	1.0	1.4	1.2	2.0	-0.2	0.2	

按照个人总分高低排列后，就要计算得分最高的25%的人（20人×25%＝5人）和得分最低的25%的人在每一个项目



上的平均分。

例如，第一个项目：

总分高的5个人平均分为：（5+4+5+4+5）/5=4.6

总分低的5人的平均分为：（2+2+1+1+1）/5=1.4

两者相减（4.6-1.4=3.2）即为第一个项目的分辨力系数。

分辨力系数越小，说明这个项目的分辨力越低，这种项目就应该删除。例如，从表6-1中可以看出，第11、12个项目的分辨力系数分别为-0.2和0.2，它们的分辨力系数是最低的，因此要删除这两道题。

2．量表的评价

完成以上步骤后，保留下来的量表从严格意义上说，也不是最终的量表，使用这个量表测出来的数值是否可靠、准确，是否有适用性，还需要对量表的信度、效度进行评价。

信度是指如果重复测量，量表所测结果的一致程度。效度指的是观测得分的变差所能反映真实得分变差的程度。对信度、效度的评价可采用第二章第五节介绍的方法。

3．量表设计中的其他注意事项

（1）量表要尽量简单，使被调查者能够尽快较容易地完成，减少由于不答而产生的偏差和提高精度。

（2）量表中项目的排列顺序将可能对调查结果产生影响，因此，量表中一半正向问题，一半负向问题要随机排列，如果可能的话，可以以不同的排列顺序制作几个版本的问卷。在电话调查和面访调查的方式下，可以通过控制访问员随机“轮换”读出项目的顺序来实现，即在问卷中访问员指导语中，表明需要轮换项目的顺序。

（3）在设计量表时，要认真考虑中性回答的问题。如果被

调查者真的是没有什么偏好，那么他们会对有双数个选择的量表（例如四级量表）中所隐含的强迫选择感到不满。如果偏好没有那么强烈，那么采用奇数个选择的量表（例如五级量表）就有可能导致回答集中在中间点上。因此在设计时，要充分考虑这两种情况。

第四节　电话调查问卷的设计

电话访问是调查研究的一种方式，这种访问方式，是新闻机构目前最常采用的调查访问方式，这是由精确新闻报道的特点所决定的。由于面对面访问需要较高的成本，而且执行调查需要时间，邮寄调查执行起来更是旷日持久，这两种调查方式往往无法在时效上满足新闻工作的要求，因此只能用于研究没有时效性的题材。遇到突发性的新闻事件、需要立刻进行调查时，比如，2001 年 7 月 13 日，北京成功获得了 2008 年奥运会的主办权，如果某新闻机构想了解北京市民听到这个消息后的反映，那么在这种情况下，电话访问是最快、最有效的信息收集方式。电话调查迅速发展也得益于电话调查设施的发展与完备，目前很多专业研究公司，以及 Call Center（呼叫中心）公司都安装了计算机辅助电话调查系统（CATI 系统）为客户提供电话调查的服务，借助于这种系统，可以很快地完成大样本的调查。因此，电话访问在精确新闻报道上扮演着越来越重要的角色。

一、用于电话调查的问卷特点

我们在前几节所讲的设计问卷的步骤和一些基本原则，对于电话调查问卷也是同样适用的。设计用于电话调查的问卷



时，也要经过确定收集信息，确定每个问题的内容、措辞和结构，确定问题顺序，确定问卷排版，对问卷进行测试等过程。但是作为电话调查适用的问卷，也有一些独特的特点。

首先作为电话调查，是由访问员通过电话与被调查者交谈并进行访问，所以不管是传统的电话调查，还是 CATI，问卷都要设计成对话式的。

其次电话访问虽然完成率比较高，但被调查者在访问中，可以轻易地拒绝回答某些问题，或是在访问过程中挂断电话，拒绝继续接受访问。所以电话调查的问卷应能够一直激发被调查者参与的兴趣，访问的问题不能太长，也不能太复杂，访问最好能在 15 分钟内完成。

电话调查的问卷不能使用图片、照片、实物等视觉性辅助资料。

在使用 CATI 方式访问的问卷中，由于使用计算机辅助访问，所以可以实现复杂的跳答模式，同时可以随机安排问题，以减小由于题目顺序造成的偏差。

二、电话调查问卷的封面信

和所有其他的调查方式一样，电话调查问卷的开始要有一封封面信，用于向被调查者说明调查的目的与重要性，调查的执行机构，以及访问所需的时间；在封面信中还要强调，该项调查只会公布整体研究结果，而被调查者的个人资料，将会被保密等等，以争取被调查者的合作。

封面信写得好，可以有效地提高访问成功率。电话访问的简介应该简短扼要，由访问员根据封面信向被调查者说明。

封面信的内容，应包括以下信息：[1]

（1）核对访问员所拨的电话号码是否正确；

（2）说明访问员的身份及调查执行机构；

（3）简单说明调查的目的以及抽样地区；

（4）促使被调查者合作的话。

下面是两个电话调查封面信的例子：

例1：您好！请问您的电话号码是不是(插入访问员所需访问的电话号码)？我是××报调查中心的访问员，我们正在进行一项民意调查，想问您几个有关北京市交通状况的问题，大概只需要耽误您10分钟的时间，请问您是不是可以接受我们的访问？

例2：您好！请问您的电话号码是不是(插入访问员所需访问的电话号码)？我是××调查公司的访问员，我们受××电视台的委托，正在进行一项有关北京申办2008年奥运会的民意调查，能不能请您花10分钟的时间，接受我们的访问？

三、电话调查问卷的题型和量表设计

在设计用于电话调查的题目和量表时，要充分考虑到电话调查的特点。前面我们介绍的题型和量表有的适合于电话调查，有些则不然。

1. 电话调查中问卷的题型设计

（1）二项选择题

二项选择题是封闭式问题中最简单的一种，适用于电话调查，可以用来询问一些事实性问题，或简单的态度问题（见第三节例）；也可以用做过滤问题。例如，某报想了解北京市居

① 罗文辉《精确新闻报道》，正中书局，1997年，第150页。



民对于改建后的王府井的意见，对于这样的调查，没有去过改建后王府井的人，就不适合接受访问。所以在问卷设计中，就有一道过滤性题目：

请问您是否去过改建后的王府井？

1．去过　　　　　2．没有去过

通过这个简单的二项选择题，可以过滤出真正想要访问的合格的被调查者。

（2）多项选择题

多项选择题是所有类型调查最常使用的题型。当然也适用于电话调查。对于多项选择题的设计，除了我们在前面第三节中提出的几点原则之外，对于需要"限选"要求的题目，例如，指导语中有"限选3项"，"只选1项"等，在设计题目回答项的时候，类别应尽量少。因为在这种情况下，被调查者往往需要听完所有可供选择的回答项，才可能从中进行选择，而电话调查，被调查者只是和访问员在电话上进行交流，而无法在访问中使用卡片等辅助访问工具，如果问卷中问题设置的回答项太多，被调查者根本记不住所有的可供选择的回答项。

（3）顺位题（排序题）

由于顺位题要求被调查者将全部或部分答案项进行排序。如下例：

下面是一些职业的名称，请您依照这些职业在您个人心目中的收入水平加以排列，请您在您认为收入水平最高的职业旁边的横线上填1，在您认为收入水平第二高的职业旁填2……依此类推，请一直填到您认为收入水平最低的职业为止。

1．警察 ________　　　　2．教师 ________

3．记者 ________　　　　4．商人 ________

5．医生 ________　　　　6．律师 ________

7. 公务员________　　　　8. 运动员________

9. 工程师________　　　　10. 工人________

这个例子要求被调查者对所有列出的回答项进行排序，很显然，这不适合电话调查，首先是题目的问法、语气，适合于被调查者自填问卷，另外对10个回答项排序，被调查者很难在电话访问的情况下记忆所有回答项，并将之排序。如果我们采用电话调查的方式，上面的例子可以采用这样的方式提问：

下面是一些职业的名称，请您依照这些职业在您个人心目中的收入水平，选出3个您认为收入水平最高的职业。

1. 警察　　　　2. 教师

3. 记者　　　　4. 商人

5. 医生　　　　6. 律师

7. 公务员　　　　8. 运动员

9. 工程师　　　　10. 工人

收入水平最高的职业________________

收入水平第二高的职业________________

收入水平第三高的职业________________

虽然我们改变了提问方式，可是在电话调查中，还需注意回答项一定不能多（不要超过10项），以免被调查者无法记忆，而难以排序。

2. 电话调查中的常见量表设计

在本章第三节中，我们介绍了几种常见量表，在电话调查中，这些量表是否适用呢？

(1) 配对比较量表

由于配对比较量表只是在多种可能的选择之间进行简单的两两比较，所以可以用于电话调查当中。

关于使用配对比较量表应注意的问题，详见本章第二节。



(2) 排序量表

排序量表要求被调查者将所有回答项中的每一项都相对于其他项目排上顺序。基于和顺位题目相同的理由，排序量表并不太适用在电话调查中使用。因为除非只有特别简单的几个项目，被调查者将很难记忆所有项目，并将这些项目排出各自相对的顺序。如果在研究中一定需要知道各个项目在被调查者心目中的相对位置，我们可以采用其他的方式实现。例如，可以要求给各个项目打分，然后在分析数据时，从各个项目的得分排出它们的相对位置。

例如：请您按照您心目中的重要程度对决定个人收入高低的因素进行排序。最重要的因素排为“1”，第二重要的因素排为“2”，依次类推，直到将所有因素都排了顺序。（请将序号填在下面的空格上）

1. 工龄　　　　____________

2. 学历　　　　____________

3. 岗位职务　　____________

4. 工作态度　　____________

5. 业务技术能力　____________

6. ……

上面这个例子，在电话调查中，就可以考虑改为下面这种提问的方式：

请您按照您心目中的重要程度对决定个人收入高低的因素进行打分。其中5分表示“非常重要”，1分表示“非常不重要”，从5到1，重要程度递减，您认为_______的重要程度得分是：

1. 工龄　　　　____________

2. 学历　　　　____________

3. 岗位职务 ________________

4. 工作态度 ________________

5. 业务技术能力 ________________

6. ……

这种打分的方式，不仅可以让我们获得项目间的相对顺序，而且可以得到项目的绝对分值和项目间的距离。

(3) 李克量表

李克量表适用于电话调查当中，有关设计李克量表的注意事项详见本章第三节。

(4) 语意差别量表

语意差别量表是用成对的反义词形容词，来测试被调查者对某个项目的态度，由于在对这个研究对象的测试过程中，每一对形容词都是从不同的维度来描述研究对象，各对形容词都是不同的，因此量表比较复杂，如果希望在电话调查中应用语意差别量表，一定注意在问题中，要清楚地告诉被调查者他们评分的确切对象是什么，评价的标准是什么，每次评价的角度是什么，也就是要让他们清楚了解每对形容词对于评价对象的意义。

(5) 斯塔普量表

我们知道，使用斯塔普量表可以生成一个轮廓，用连续数据说明量表中的每个词汇或短语对所研究的对象描述了多少。在电话调查中使用斯塔普量表也可以包含比较大量的词汇或短语，要求被调查者进行评价。但和语意差别量表一样，将斯塔普量表用于电话调查时，也要注意让被调查者明确他们所要评价的对象是什么，而打分或判断的标准又是什么。



四、计算机辅助电话调查（CATI）的问卷设计

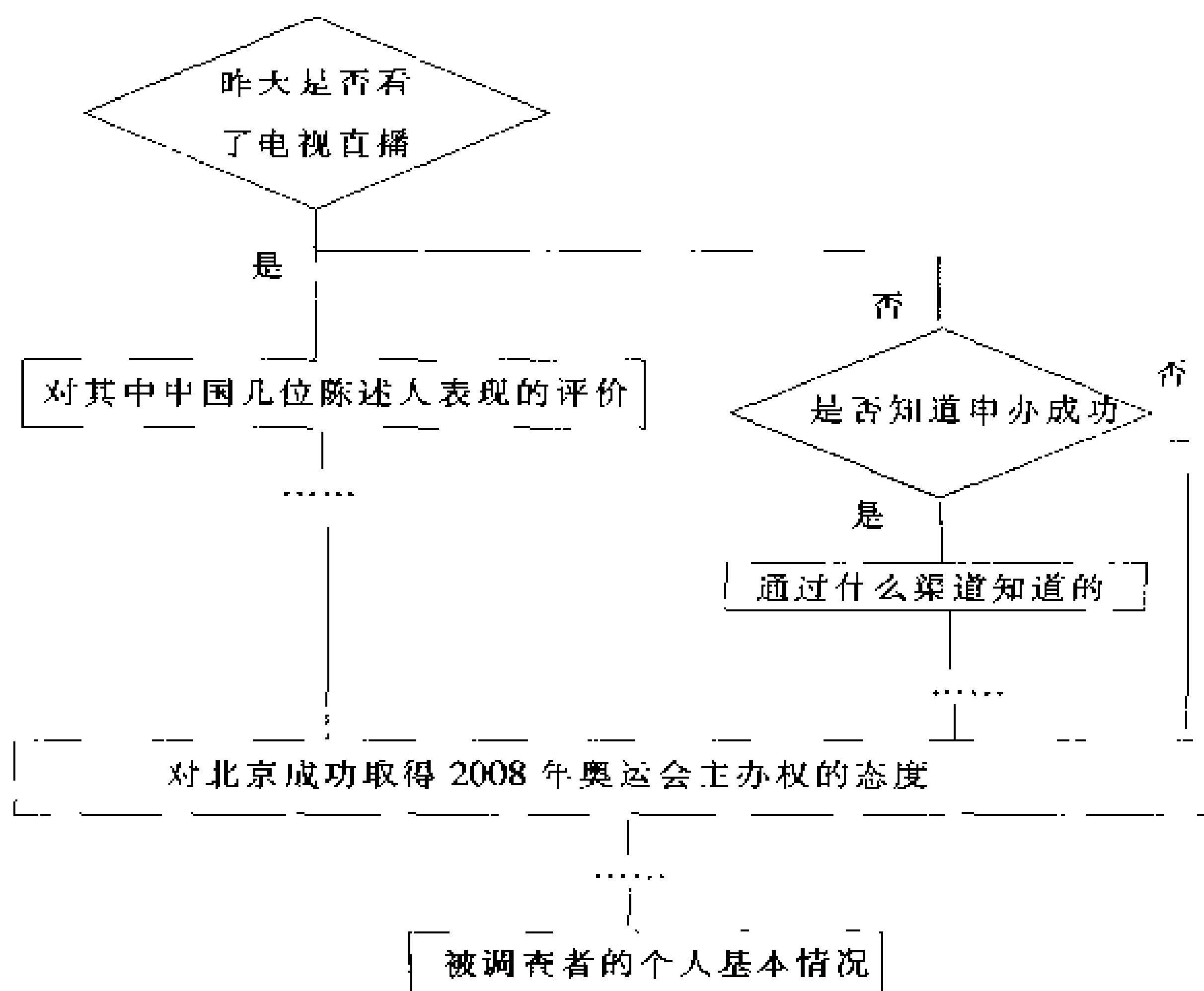
随着计算机技术的发展，计算机辅助电话调查的方式渐渐取代传统的电话调查，一般的 CATI 软件中，都带有问卷生成系统。使用这种系统，可以很方便地完成问卷设计的工作。在利用计算机问卷生成系统设计问卷时，除了要注意我们在前面提到的问题之外，也要充分利用这种系统的独特的优点。

首先在 CATI 方式下，借助于计算机的自动控制，可以实现复杂的跳答模式。因此我们在设计问卷时，可以不必顾虑复杂的跳答模式可能导致的访问员访问误差，但是在设计这些跳答模式时，最好使用我们前面介绍的框图设计法。即在问卷设计之初，先画出整个问卷的各个部分及前后顺序的框图，并按照问卷的逻辑结构，标明分叉问题及跳答的模式，然后再具体写出每一部分中的具体问题，并安排好这些题目间的相互顺序。这样我们问卷生成系统中设计问卷时，就可以明确地告诉计算机问卷中各个题目间的相关性，从而利用计算机的自动控制程序，在访问时，实现跳答模式的正确控制。下面是一个简单的框图例子。

某报在 2001 年 7 月 14 日，也就是北京申办 2008 奥运会成功的第二天，希望进行一项民意调查，了解北京市民的态度。问卷的结构见第 272 页图：

除了可以借助计算机实现复杂跳答模式之外，在 CATI 方式下，还可以通过计算机实现回答项及量表项目顺序的随机化。有研究表明，问卷中问题回答项和量表中项目的排列顺序对调查结果将产生影响，为了消除这种顺序偏差，在其他的访问方式下，可能需要制作不同的问卷版本或控制访问员轮换回答项或量表项目的顺序，但这两种方式都不是最好的，制作不

同版本的问卷有很多时候是不现实、不经济的；而对于访问员，并不能实现完全控制他们的访问顺序。在 CATI 方式下，情况就变得很简单。借助计算机控制程序，可以使得访问员每次访问使用的问卷，其问题的回答项或量表中项目的顺序是随机的，这样就可以很好地消除顺序偏差。





第九章　调查数据的整理

通过调查收集的数据，要进行适当的整理，这个过程的目的是将数据转换为适合进行数据分析的形式，包括数据收集完毕之后、估计之前对数据进行加工处理的所有工作。数据整理的整项工作费时、费力，而且需要非常仔细、认真，是一个繁琐的事情。但是数据整理对数据最终的质量和成本却有着很大的影响。因此，需要对数据整理工作进行周密的计划，同时实施质量监控，严把数据整理的每一关，并在必要时进行校正。

数据处理的整个过程既有自动完成的，也有手工完成的，这与调查采用的数据收集方法有关。以一份最基本的印刷在纸张上的问卷为例，调查数据的整理过程大致包括以下几个部分：

（1）调查问卷的接收、检查和校订；

（2）数据编码和录入；

（3）数据审核；

（4）对数据的统计预处理。

精确新闻报道调查经常在调查实施部分采用计算机辅助电话访问（CATI），如果调查采用 CATI 方式或计算机辅助面访（CAPI），那么上面所说的调查数据的整理过程中，数据录入以

及之前的工作是与数据收集同时完成的，这样可以省去很多前期的、繁琐的手工操作，也减少了错误产生的环节。

第一节　调查资料的接收、检查和校订

调查数据的整理工作是从现场实施回收第一份问卷开始的。

一、接收与核对问卷

对于一项调查研究的问卷，要认真细致地做好接收和核对工作。

1. 接收问卷

负责接收问卷的人员要事先设计好一定的接收文件系统，比如，接收表格等等，用于登记自己收到的问卷，如果是面访调查，这些表格的内容可能包括访问员姓名、调查地区、调查实施的时间、交付的时间、实发问卷数、上交问卷数、未答或拒答的数目以及合格问卷数等；如果是邮寄调查，可根据实际情况设计相应的问卷接收表格。接收人员要掌握每天完成的问卷数和每天接收到的问卷数。

接收人员要在收到的问卷上注明接收日期，以便日后有需要时，有据可查。

回收的问卷就是研究的原始文件。对每一份回收的问卷都要记录一个惟一的、有顺序的识别号码，如果有几个接收人员同时工作，那么最好事先分别号码的范围，比如，接收员A对接收收到的问卷编号时，范围从001～050，接收员B对接收到的问卷编号时，范围则规定为051～100，等等。这些惟一的编号不但记录在原始问卷中，而且也将同时录入到数据中。



在对数据进行差错，或有其他需要核对原始问卷时，研究者可以随时找到和数据对应的原始文件。

对收到的原始问卷要认真保存，刚刚回收的问卷可能要由不同的人员去进行核对、编码、录入等后期的工作，在这个过程中，要注意按照问卷的识别号码及时登记，记录何人拿着哪些原始问卷。建立系统的登记制度是非常必要的。

作为研究者，要仔细控制数据收集和整理的过程，掌握实施和问卷回收、整理的整个进度。

2．核对问卷

在接收问卷的同时，要对问卷进行粗略的核查，检查问卷的完整性和实施质量，确定哪些问卷可以接受，哪些需要作废。

在进行核查之前，首先要制定核查的规则，告诉所有参与核查的人员什么样的问卷是可接受的，以保证所有人都按照相同的规则来检查回收的问卷。对每份回收的问卷都要做彻底的检查，包括每一页和每个部分，查看是否填写正确，是否符合规则。通过检查，应该将所有原始问卷分为3个部分。即：

(1) 可以接受的问卷；

(2) 明显需要作废的问卷；

(3) 对是不是接受有疑问的问卷。

例如，下面几种情况的问卷一般是不能接受的：

缺损的问卷。比如，缺了一页或数页的问卷。

回答不完全的问卷。比如，相当多的内容没有回答。

从问卷的填答情况来看，访问员（被调查者）并没有理解或遵循访问（回答）指南，或没有理解问题的内容而错误填答的。

问卷中问题的答案异常，比如，在使用李克量表测量的题

目中，对所有项目的问答几乎是一样的，等等。

被调查者是不符合调查要求的人，例如，一项调查要求被调查者是具有法定选举资格的人，而其中一份问卷的被调查者年龄为17岁，这显然不符合要求。

核查完成后，可以接受的问卷就进行编号，准备进入后期工作，而不符合要求的问卷要作废，核查人员有疑问的问卷应由研究人员决定取舍。

如果对所研究的样本有配额的要求，或对某些子样本有具体数量上的规定，那么此时核查人员就应该分类清点合格问卷，如果在数量上没有满足相应的要求，可能就要根据需要对样本做一些补充。

二、检查与校订问卷

1. 对问卷的进一步检查

这时候对问卷所做的检查，要比上一步更加细致、严格，主要检查问卷填写得是否清楚可认，是否填答完全，是否一致，是否正确等等。在检查中要特别关注以下问题：

(1) 如果问卷中含有开放题，特别是开放题数量较多时，不论是访问员记录的答案，还是被调查者自己填写的答案，都有可能字迹模糊，无法辨认。

(2) 如果问卷中有个别的问题没有回答，那么就会出现答案不完整的情况。

(3) 有些时候，可能问卷中的答案出现明显不一致的情况。例如，有一个18岁的被调查者，其婚姻状况是已婚。

(4) 没有按照调查指导语的要求回答的，比如，单选问题，而填了多个答案，或者分叉问题出现错误的。

例如：Q1 您吸烟吗?



1．是，吸烟

2．不吸烟……跳答 Q5

对于有这样分叉要求的题目，检查人员要认真检查，保证问卷的逻辑结构正确。

含有以上几种情况的问卷，都是有问题的问卷，要做进一步的处理。

如果有几个检查人员同时做这项工作，在工作开始之初，要将需要重点检查的题目告诉每个人，比如，开放题含有分叉要求的问题、可能会出现前后不一致的问题，等等。

2．对有问题问卷的校订和处理

在进一步检查过程中发现有问题的问卷，要进行校订和处理。比如，在含有分叉要求的题目中，被调查者填写了本不该自己回答的项目，检查人员就要对此做必要修改；对于字迹模糊的开放题，可以辨认的，要在原来的答案旁边以清楚的字迹重新注明，等等。对于不能通过校订解决问题的问卷，可以考虑以下 3 种处理办法：

(1) 退回实施现场，要求访问员对被调查者进行回访，以此来获得比较满意的数据。

(2) 对不符合要求的问题按照缺失值处理，有的时候，对被调查者进行回访是不现实的。比如，不能确认被调查者，或时间与经费的限制，不能进行回访等，而且有不满意答案的问卷比例很小，不满意答案在整份问卷中的比例也很小，涉及的又不是关键变量。在这种情况下，可以考虑将不满意的答案按照缺失值进行处理。

(3) 作废整份问卷。如果不满足按照缺失值处理的条件，比如，不满意的答案在整份问卷中占的比例很大，涉及的又是比较关键的变量，而整个研究的样本量很大，有问题的问卷所

占比例很小，这时候，可以采取将有问题的问卷简单作废的处理方法。

第二节 调查数据的编码和录入

对调查数据编码就是确定问卷中各问题答案所对应的计算机代码（一般是数字），便于数据录入和处理。编码方案的制定，应该以能将调查中所得到的数据分成有意义的类别并得出各类别的基本模式为指导方针。编码时最好是先列一个不能再细分的大表，然后在此表格的基础上再根据具体需要对某些类别进行合并。因为分类太少会掩盖某些信息，产生误导，而且，“其他”类太大对研究问题也没有益处。理想的情况是类别列表既短又全面，且选项之间没有重叠。但这往往很难做到。

编码可以分为两类，即事前编码和事后编码。事前编码在设计问卷时进行，事后编码在数据收集完成后进行。

一、事前编码（Pre-coding）

对于结构式问卷中的封闭式问题，其回答项是在设计问卷时就已经确定的，可以保证大多数的答案都会落入事先确定的类别中。事先编码时，每个回答项所对应的数字代码通常会印在问卷上，有时甚至数据文件的记录格式也在问卷中体现出来。见下例：[1]

在您所关注的不同的能源问题中，下面哪一项是您最关

[1] 资料来源：北京广播学院调查统计研究所 1998 年《IEM30 国世界环境意识调查》项目。



注的？

（读出，只选一项）　　[12-13]

1．可能的能源短缺

2．核能产生的风险

3．能源生产和运输对自然的损害

4．使用矿物燃料，例如煤和石油（汽油）对地球气候的影响

5．可能的能源价格增长

自愿的回答（不要读）

6．不关心

7．其他（请详述）________

9．不知道/无回答

以上这个例子中，回答项旁边的数字即为回答项的编码，题目下面方框中的数字即为数据文件中的该问题对应变量的列位置。当然很多时候，这个列位置在问卷上是不出现的，因为我们可以在数据录入时再规定变量的列位置。

对问卷进行事前编码可以节省大量的时间和费用，但是事前编码也具有一定的局限性，我们没法确切知道被调查者原始的回答，这就会给验证编码的质量带来困难，而且问卷回收后也没法更改编码。比如，某报进行一项关于市民在如何安排业余时间的调查，在问卷设计期间事先编码时，将业余时间可能进行的活动分为休闲、学习提高等等。数据收集回来后，发现市民在业余时间进行休闲活动的比例最高，而休闲这个概念所涵盖的范围太大，记者们想进一步知道被调查者所从事的是什么样的休闲活动，但是由于事前编码太过粗略，数据回收后已经没有办法得到再细致的信息了。

在制定事前编码时，也要考虑到调查所采用的方式，在邮

寄或面访调查中，我们可以将类别设计得多一些，比如 20 个，让被调查者从中进行选择；而如果是电话调查，在电话中由访问员将 20 个类别念给被调查者听，是很困难的。

二、事后编码

在数据收集回来之后，如果某个问题没有事先设置编码，就需要进行事后编码。通常需要事后编码的情况主要有：

1. 混合式问题中的“其他”回答项

在混合式问题中，如果问题所列出的回答项没有包括被调查者的回答，那么这一回答在实施过程中将归入“其他”类别，在要求对“其他”类别进行详细说明的情况下，这样的答案就没有事先规定的代码，因此在数据录入之前要先完成事后编码的工作。

2. 开放式问题

对于开放式问题，一般都需要进行事后编码的工作，而且对于开放问题的编码，工作难度更大。

目前进行事后编码的方式主要有两种，即手工编码和自动编码。事后编码同样要遵循编码的指导方针。

(1) 手工编码

手工编码过程中，编码员要阅读、理解问题的书面回答并把它归类，编成相应的数字代码，数字代码可以直接录入计算机，也可以标在问卷上。编码可能基于对一个问题的回答；也可能要基于对几个相关问题的回答，这时，编码的质量在很大程度上就取决于编码方案的完善性和编码员的技能。

在进行编码操作时，可以发给每个编码员一份空白问卷，将需要考察和编码的问题明确地标在问卷上，以避免遗漏。

每个需要编码的问题都要有一个编码表，表上除该问题已



经确定的编码外，还要清楚标示问题的题号，并预留足够的空间来添加新码。

这样的编码表要给每个编码员人手一份，都是完全相同的，所有编码员要集中在同一地点工作。因为如果编码员不在同一地点工作，他们就无法知道其他编码员在编码表上设立了什么新码，因此会导致同一代码对应不同的答案。

对于编码员的工作，研究人员要规定尽可能具体的编码准则，告诉编码员如何识别答案、如何将其归入一定的类别内，等等。例如，"沙尘"是否应该归入"空气污染"一类。

编码员要进行严格的培训，因为在编码时需要考虑以下问题：

①可能的答案数；

②对答案进行判断的复杂性；

③回答的不明确性（即回答的质量不高）。

编码人员之间的差别是在所难免的，因此，对每一位编码员所做的第一批问卷都应该逐份进行仔细的检查，找出错误，并且确认是否有必要进行进一步的培训。之后，还要对编码质量进行周期性的检查并且在必要时加以纠正。

在编码过程中，要遵循这样的原则，设立较多较窄的类别要优于设立较少、较宽的类别。当编码员很难决定是再设立一个新代码，还是将其合并到已有的一类中去，这个时候，一般要选择设立一个新代码。因为如果需要合并，可以在数据分析时通过计算机很容易的实现。反之则不然，一旦在编码时已经合并了，在数据分析时就没有办法再区分，除非去查找原始的问卷。

（2）自动编码

所谓自动编码是通过计算机文本识别软件自动编码（即用

软件给一串字符分配一个代码)。多用于英文系统中。

三、编码说明

在完成事后编码之后，问卷中所有的问题都被编了码，在进行数据录入之前，我们有必要将问卷中所有编码信息，包括事前编码和事后编码，集中到一起，编写一份编码说明。以便于了解变量编码信息和建立数据录入结构。一般的编码说明应该主要包含以下几个内容：

(1) 变量编号；

(2) 变量名称；

(3) 变量标签 (Variable Label)；

(4) 变量所在的列位置；

(5) 变量的编码说明 (变取值标签 (Value Label))。例如表9-1：

表 9-1　　北京市民对申奥态度的调查编码说明

题目编号	变量名称	变量标签	列位置	变量编码说明
	No	问卷编号	1~3	001~600
1	Q1	对北京申奥的关注程度	4	1. 非常关注 2. 比较关注 3. 说不清 4. 不太关注 5. 根本不关注 9. 不知道/无回答

续表

题目编号	变量名称	变量标签	列位置	变量编码说明
2	Q21	城市－大阪	5	0．未答　1．大阪
	Q22	城市－巴黎	6	0．未答　1．巴黎
	Q23	城市-伊斯坦布尔	7	0．未答　1．伊斯坦布尔
	Q24	城市-开普顿	8	0．未答　1．开普顿
	Q25	城市-多伦多	9	0．未答　1．多伦多
……				
32	Q11	被调查者性别	46	1．男　　2．女
33	Q12	被调查者的文化程度	47	1．小学及以下 2．初中 3．高中（中专，技校） 4．大专 5．本科及以上 9．无回答
……				

从这个例子可以看出，编码说明提供了有关变量的定义、取值、编码意义等方面的内容。

四、调查数据的录入

1．*提高数据录入效率的方式*

单纯的做数据录入工作是一项单调，而且容易出错的工作，如果是大规模的调查，数据录入也是一项很花时间的工作。如果要提高数据录入的效率，可以采用以下几种方法：

（1）采用计算机辅助数据收集方法；



（2）对纸张式问卷进行光电扫描；

（3）对纸张式问卷进行事先编码。

如果在调查中采用的是CATI或CAPI的方式，那么在问卷生成的同时，录入数据库的结构也应该完成，访问员在访问的同时，也就将被调查者的回答，也就是收集到的数据录入到数据库中了，所以数据录入是和数据收集的工作同时完成的。

而如果采用传统的调查方式，数据录入是在数据收集完毕之后（通常是在对问卷进行初步的审核和编码之后）进行。光电扫描就可以大大提高数据录入的速度，问卷较短时效果更加明显。与手工录入相比，光电扫描可以减少数据录入的差错。

但是就我国目前的情况来看，手工键盘录入还是最主要采用的录入方式。

2．设计数据库结构

为了进行数据录入，首先要建立录入的数据库结构，有很多软件可以实现数据录入的功能，如数据库软件Foxbase、Foxpro等；统计软件SPSS、SAS等；大家非常熟悉的办公软件Excel也可以用来进行数据录入；当然还有专用的数据录入软件，如基于DOS环境的Pcedit等软件。

在设计数据库结构时，要注意以下问题：

（1）规定变量

调查问卷中的每个问题都要用一个或多个变量来对应，例如在表9-1中，我们用Q11来代表被调查者的性别，Q11＝1，代表男性；Q11＝2，代表女性。

在规定变量的同时，还要规定变量的类型，变量的取值范围，变量的位数，其中小数点所占的位数，等等。变量的类型常用的有字符型和数值型两种。字符型的变量可以输入任何字符，不一定是数字，但是对字符型的变量，除了进行频数统



计，不能做其他任何数值性的计算。所以更为常用的还是数值型的变量。例如，对于性别 Q11，规定为数值型变量，取值为 1 和 2，这样将来可以运用多种统计方法进行分析。

规定变量的取值范围时，对于尚未完成事后编码的问题，要注意留出足够的空间。另外还要考虑到数据缺失的情况，这时可以考虑用一些特殊的数值（例如，“0”，“9”，“99”）来标示这些缺失值。如，表 9-1 中的变量 Q12：

Q12 ＝1．小学及以下

＝2．初中

＝3．高中（中专，技校）

＝4．大专

＝5．本科及以上

＝9．无回答

变量的位数是指录入数据时，一个值最多占几位。例如，表 9-1 中，NO（问卷编号）这一变量，要占 3 位，而因为这个变量的取值均为整数，所以它的小数点位数为零。

（2）确定问题对应的变量数

①单选问题

对于每一个单选问题，只要规定一个变量与之对应就可以了。例如，表 9-1 中的 Q1 题：

Q1. 请您告诉我您对北京申办 2008 年奥运会的关注程度如何，您说您是非常关注，比较关注，说不清，不太关注，还是根本不关注？（只选一项）

1．非常关注　2．比较关注　　3．说不清

4．不太关注　5．根本不关注

这个问题是单选问题，所以我们在编码说明中，只用一个变量 Q1 来代表。

②多选问题

多选问题要根据可以选择的选项数目来决定与之对应的变量数，一般来说，变量的个数要等于可供选择的答案个数。例如，表 9-1 中的 Q2 题：

Q2. 请问您知道以下城市，哪些是本次和北京一起申办 2008 年奥运会主办权的城市吗？（可多选）

1. 大阪　　2. 巴黎　　3. 伊斯坦布尔

4. 开普顿　　5. 多伦多

对于这样一个问题，就需要用 5 个变量来对应：

Q21 = 1. 大阪　　Q22 = 1. 巴黎　　Q23 = 1. 伊斯坦布尔
　　= 0. 未答　　　　= 0. 未答　　　　= 0. 未答

Q24 = 1. 开普顿　　Q25 = 1. 多伦多
　　= 0. 未答　　　　= 0. 未答

③开放式问题

对于开放题目录入的处理有几种方式，如果不打算对开放题进行定量分析，那么就没有必要进行编码，在录入时，只要简单地设计一个 0—1 变量，其中可以用“1”表示回答了开放问题，“0”表示没有回答；如果计划对开放题目做定量的统计分析，在录入之前，就要对开放题进行编码，再根据是单选问题，还是多选问题来规定变量与之对应。

④跳答模式

很多问卷有跳答模式的设计，在建立录入数据库时，要对这个问题加以考虑，有的软件可以实现对跳答的控制，在设计数据库时，就要充分利用这些功能，可以使得录入更加方便，而且准确性更高。

3. 数据核查

如果采用人工键盘录入的方式录入原始数据，那么无论如



何控制，差错也还是有可能发生，因此要对录入工作进行核查。

全面的核查要求双机录入，即每个记录都要录两次，然后对两次录入的数据逐个比较，如果录的不同，错误就被检测出来了。但是对整个数据集进行全面核查，时间和费用都要加倍。因此出于对时间和费用的考虑，研究人员通常会采用部分复查的方式，来检查录入的质量，一般来说，抽查其中 25% 左右就可以了。如果只发现少量错误，那么改正错误即可；如果发现大量的错误，可能就要考虑重新录入的问题了。

第三节　数据审核

数据录入完成之后，理想情况下，每一条数据都应该是完整而且没有差错的，但是各种原因可能会导致数据出现问题。比如，前期对问卷的检查不够细致，问卷中尚有一些错误，或者录入人员的录入错误等。因此在数据录入后，要对数据进行审核，审核的目的就是要保证调查最后得到数据的完整性、一致性和有效性。

在这个阶段对数据进行审核是非常重要的，如果数据中存在问题，可能会无法进行下一步的数据分析，或者数据分析已经完成，对调查结果报道也已经呈现给受众，而研究人员并没有意识到数据里面的许多错误。

一、审核的种类及审核规则的确定

1. 审核的种类

审核通常可分为三类：有效性审核、一致性审核、分布审核。

（1）有效性审核

有效性审核主要检查问卷上的回答是否是有效的。如，被调查者（访问员）是否按照问卷上的调查指导语完成问卷，是否在要填数字的地方填上了非数字字符，数据是否在允许范围之内，问卷中的回答是否有缺失等在内的差错等。如，问卷中有一道题目是：

请问您出生的年份________________

如果研究设计要求样本的年龄在18～65岁之间，那么上面这个问题的取值范围就应该是1936～1983之间，如果发现某条数据这个答案的值是1985，那么这个数据显然不符合调查的要求。再比如，根据表9-1的编码说明，Q11题的取值只能是“1”或“2”，如果数据集中发现其他的取值，那么就是超出范围的无效数据。

（2）一致性审核

一致性审核检查问卷中不同问题之间的关系是否正确，一致性审核可以基于不同的问题或同一问题的不同部分之间的结构关系、逻辑性和合法性来进行检查。如，出生年月和职业状况，对于25周岁公民来说，如果他的职业状况填的是“退休”，这显然不和常理。另外对于问卷题目之间逻辑结构，诸如跳答模式是否正确等的检查，也属于一致性审核的范围。例如，在“您昨天是否看过报纸”这个问题中回答“否”的人，下面关于“昨天看了哪些报纸”和“昨天看了多长时间报纸”等问题就不必回答，如果回答了，那么也就会出现数据不一致的错误。

（3）分布审核

分布审核主要是借助计算机拟和数据的统计分布，确认异常记录。分布审核有时也称为统计审核或异常值探测。



所谓异常值是指和数据库中其他数据不一致的观测数据。异常值分为单变量异常值和多变量异常值。如果一个异常值对应一个变量，这就是一个单变量异常值；否则就是多变量异常值。例如，某个家庭人口数是10人，而家庭月收入为500元，这就是一个多变量异常值。

异常值的出现可能有两种原因：第一是数据有错误，如数据录入错误；第二是数据固有的多变性导致异常值的出现。

比如，在录人的数据中，发现某人的月收入是10万元人民币，这个数据的出现可能是由于录入员录错了，也可能这个人的收人就是10万元。

通常的做法，异常值是通过测量它们和数据中心的相对距离来辨认的。例如，如果 y_1，Y_2，Λ，y_n 是要观测的样本对应某一个变量的数据，m 和 s 分别是测量该变量数据集中趋势和离散趋势的指标．那么，Y_i 离数据中心的相对距离 d_i 就可以定义为：

$$d_i=\frac{|y_i-m|}{s}$$

如果这个距离 d_i 越出了预先确定的偏离值，那么该观测数据就被认为是异常值。

一般情况下，样本均值和样本方差是用来测度数据集中趋势和离散趋势最常用的统计量。但是，由于它们对异常值比较敏感，因此，选择它们就不太合适。例如，如果数据成群偏在一边，样本均值就会偏向异常值，样本方差也会由于异常值而大大放大。因此，根据有些异常值计算的 d 值就会显得相当小，确认这些异常值就较为困难。

所以，很多时候，研究者会放弃对异常值敏感的平均值和方差这两个统计量，而采用四分位数法。这种方法用中位数测

度数据的集中趋势，四分位域测度数据的离散程度，因为这些统计量对异常值不大敏感。

严格说来，数据审核贯穿于整个调查过程，从访问员在调查实施时所做的简单的、初步的检查，问卷回收后所做的检查和校订，到数据处理阶段数据录入之后由计算机程序进行的复杂的校验，都可以称为审核。一般将访问员收集数据过程中所做的审核叫做现场审核，现场审核一般包括对所收集到的数据进行有效性审核，以及一些简单的一致性审核。同样在数据录入之前，对回收的问卷所做的审核也以有效性审核和一致性审核为主。较复杂的审核一般是在数据录入之后作为一个独立的环节来进行。

2．确定审核规则

在对数据进行审核的过程中，要运用各种规则来辨别缺失、无效或不一致的数据，这些数据表明记录存在潜在的错误。

在实际工作中，可以从以下几个方面着手确定审核规则：

（1）关于调查主题的专业知识

熟悉调查主题的人，或对审核类似的变量有经验的人，对变量之间的相互关系会有很好的了解，他们的知识和经验对于审核规则的确定非常有帮助。

（2）问卷的结构

问卷有其自身的逻辑结构，很多问卷还包含有跳答的模式，指示被调查者他们应该回答问卷上的哪些题目。因此可以根据这些逻辑关系来确定审核规则，以确定被调查者的回答是否符合问卷的逻辑结构。

（3）其他相关的调查或数据

比如，在一个月前通过随机抽样进行的调查研究了北京市



居民家庭的月平均收入，而这些数据在我们一个月后进行的调查中，对同样的变量就可以有很好的参考作用。

(4) 统计理论

可以依据统计理论制定审核规则，如确定什么是异常值等。

下面是审核规则的一些具体例子：

①Q11 题有且只有一个回答

②对问题 Q3，有效的回答只能是“1”或“2”

③在 Q3 题中回答“2”的人，不需回答 Q4 题

④Q7 题各个项目的回答合计应为 100%

如果在审核过程中发现数据与审核规则不符，说明数据出现问题。

二、审核的方法

在进行审核时，审核人员的经验是非常重要的因素，一般来说，应该由那些对本次调查的主题、问卷设计和数据分析有丰富经验的人来担任审核的工作。审核工作的目标是更好地检查数据，发现其中缺失、无效、异常或不一致的数据，删除无效记录，并分离出需要进行进一步处理的数据。

审核是一个需要认真、仔细的过程，随着计算机技术的发展，可以借助一些现成的统计软件，如 SPSS、SAS 等，来实现对数据的自动审核。寻找超出范围的无效值、异常值，或按照问卷的逻辑结构前后不一致的数据。

1. 频数统计

在对录入后的数据进行审核之初，首先要对问卷中所有的非连续变量做一份频数统计，根据这份频数表，很容易可以确认那些超出变量应有取值范围的无效值。例如，在前面第二节

表9-1中，我们知道Q12题是了解被调查者的文化程度，如果用统计软件SPSS做出的频数表见表9-2：

表9-2 利用频数表检查无效值

文化程度

	频数	百分比	有效百分比	累积百分比
1	12	6.0	6.0	6.0
2	31	15.5	15.5	21.5
3	93	46.5	46.5	68.0
4	26	13.0	13.0	81.0
5	33	16.5	16.5	97.5
6	2	1.0	1.0	98.5
9	3	1.5	1.5	100.0
合计	200	100.0	100.0	

表9-1的编码说明表明，Q12的取值范围应该为1到5，9表示缺失的数据，而在上面Q12的频数表中，出现了取值为“6”的数据，这显然是超出变量取值范围的无效数据。我们可以利用统计软件，通过这个无效值以及其他相关的定位变量，在数据库中找到这个原始记录，同时可以通过问卷编号，找到原始问卷，以查询这个无效数据的由来，并加以改正。

当然对于定距和定比变量，也可以通过频数表发现问题。例如，计算市民家庭月收入的频数表，之后发现，某一数据值为1501元，那么我们完全有理由怀疑这个数据可能是录错了，因为一般回答家庭月收入不会出现个位的尾数。

2、计算变量的统计量

对于连续型变量，可以利用统计软件很方便地计算出变量的描述统计量。如，均值、标准差、四分位数、最大值、最小值等，那么按照前面介绍的方法，就可以判断出极端值。表9-3是利用SPSS计算的某次调查中北京市民家庭月收入的描述



性统计量。

表 9-3 利用描述性统计量确定异常值

Statistics

家庭收入

N	Valid	166
	Missing	34
Mean		5214.28
Median		2000.00
Std. Deviation		20181.23
Minimum		500
Maximum		250000
Percentiles	25	1500.00
	50	2000.00
	75	3000.00

其中家庭月收入的平均值为 5214.28 元，中位数为 2000 元，第一个四分位数 $q_{.25}$ 为 1500 元，第三个四分位数 $q_{.75}$ 为 3000 元，那么可以利用前面介绍的方法，计算变量取值和数据中心的相对距离来辩认可能的异常值。比如，在上表中的最大值和最小值。

当然这只是一个例子。

很多时候，如果要求不是那么严格，研究人员并不去计算某个连续变量取值的置信区间，而更多的是凭经验、知识和相关变量去判断异常值。

3．交互分析（列联表）

对于数据集中前后不一致的数据，或没有遵循问卷逻辑关系的数据，也可以利用统计软件，通过交互分析的方法来检查。例如，我们通过 SPSS 得到下面这样一张交互分析列联表：

表 9-4　利用列联表检查不一致的数据

A1 是否看过电视剧《让爱做主》?	A2 对王志文在剧中的表演的喜欢程度					
	1. 非常喜欢	2. 比较喜欢	3. 没感觉	4. 不太喜欢	5. 非常不喜欢	9. 缺失值
1. 看过	51	45	18	32	4	
2. 没看过	3					147

在表 9-4 中，有 3 个被调查者根本“没有看过”电视剧《让爱做主》，但却“非常喜欢”王志文在该剧中的表演，这显然不合乎逻辑。那么我们通过交互分析发现了错误，同样可以利用统计软件对这 3 个错误的数据在数据集中定位，并进行相应的修改。

以上介绍了 3 种利用统计软件，借助于计算机对数据进行自动审核的方法，这 3 种方法只是比较常用的方法，当然还有其他的方法可以利用。

数据录入之后的审核工作是非常重要的，它是以改善数据质量为目的的，这个环节的工作，将使我们有可能得到具有内部一致性的数据集。

第四节　对数据的统计预处理

经过审核的数据集在进行数据分析之前，还需要进行一些前期的处理，在审核过程中发现的数据错误，有些可以在审核的同时加以定位、修改，而有些错误可能就要等到这个阶段再进行处理，比如，数据集中的异常值、缺失值等。另外，根据需要对数据集中的变量进行某些转换，或对数据集做加权处理等，都是这个阶段的工作。统计预处理并不是任何时候都会需



要，但是适当的预处理可以使数据分析的质量有所提高。

一、缺失数据的处理

缺失数据是指数据集中未知的变量值。通常情况下，在数据集中有少量的缺失数据，研究人员是可以容忍的，但缺失数据必须控制在一定的比例之内，否则可能会对分析结果产生严重的影响。对于缺失数据的处理可以参考以下3种方法：

1. 推算（Imputation）

这种推算方法的指导思想，就是寻找一个合适的值来代替缺失数据。那么如何找到这个合适的值呢？

在推算方法中，比较常用的推算途径有：

(1) 演绎推理

在这种方式下，缺失的或不一致的数据肯定能通过推理得到。这种推理常常是根据问卷上被调查者对其他相关问题的回答来进行的。例如，要了解被调查者每个月收入如何分配，在每个项目中的分配比例怎样，问卷中一共设置了5个项目，包括储蓄比例、娱乐支出比例等，那么我们知道5个项目比例加起来的总和应该是100%，如果被调查者只填了其中4项，那么我们就可以根据已填4项的值，来推算第5个项目所占的比例是多少。演绎推理的估算方法优于其他的方法，因此如果可能的话，要优先考虑这种方法。

(2) 用一个样本统计量的值代替缺失数据

采用这种推算方式，用的最多的是使用变量的平均值来代替缺失数据，因为这样做的结果，该变量的平均值不会改变，那么其他的样本统计量，比如，标准差、相关系数等也不会受到太大的影响。

使用变量的平均值来代替缺失数据，可以用所有样本在该

变量上的平均值，也可以先通过其他与缺失数据变量有显著相关的变量，建立一个子样本，然后用子样本的平均值来代替缺失数据。举一个常见的例子，某一被调查者的收入缺失，为了推算这个缺失值，研究人员认为一个人的收入是和他的职业、学历等变量相关的，因此，首先按照职业、学历变量建立一个和需要推算其收入的被调查者有相似特征的子样本，然后用这个子样本中收入的平均值来代替缺失的收入变量的值。不过再怎么样，这个数据终究是估算的，如果被调查者回答了他的收入的话，可以想像，其答案可能是高于用来代替它的平均值，当然也有可能低于这个平均值。

在利用样本统计量代替缺失数据时，也可以考虑选择其他的统计量，如，中位数等对异常值不敏感的统计量。

(3) 用从一个统计模型计算出来的值代替缺失数据

这种方法的思路是寻找含缺失数据的变量和其他与之相关变量之间的关系，以此建立一个统计模型，并由该模型输入其他已知变量的值，来推算缺失的变量值。比较常见的有回归模型、判别分析模型等。

例1：某研究中的一个变量是城市居民每个月在娱乐方面支出的费用（Y）是多少，如果这个变量中某个值缺失，而研究者认为娱乐费用这个变量与被调查者的年龄（X1）和他们的月收入（X2）有关，并认为这3个变量之间存在线性关系，由此构造一个回归模型：

$$Y = aX1 + bX2 + c$$

利用已知的样本数据，可以求出上式中的系数（a，b）和常数项c。模型构造好以后，就可以利用被调查者的年龄值和月收入来估算他（她）每个月花在娱乐上的费用。

例2：某报进行大选的选举预测，在问被调查者他们在大



选中会投谁的票时，很大比例的人表示“还没有决定”，那么在进行数据处理时，如果简单地将这一部分答案删除，预测结果肯定会出现严重偏差。为了解决这个问题，研究者采用判别分析的方法，认为被调查者的社会地位、职业、党派、教育程度、生活形态等变量会影响他们的投票决定。因此，利用这些独立变量和已知的变量值建立一个判别函数，使之能够区分那些已经决定投票给不同候选人的群体。而对于“还没有决定”的被调查者，只要将他们在各个独立变量上的取值代入判别函数，就可以通过函数加以推算他（她）会投哪个候选人的票。通过建立这样的一个判别函数，可以大大提高选举预测的成功率。

使用统计模型进行推算得到的值比简单地利用样本平均值来代替缺式数据要更加稳定。统计模型可以保持相关变量之间的关系，但在使用的时候要注意，不要人为地造成数据之间的关系。

2．个案删除（List wise）

个案删除是将有缺失数据的个案整个删除的方法[①]。由于许多被调查者回答都多多少少地会有一些变量是缺失的，如果将有缺失数据的个案全部删除，其结果肯定导致样本量很小。而研究者花费时间和成本去收集数据，也不会希望大量地删除数据。而且，如果被删除的被调查者与回答了全部问题而被保留在数据集中的被调查者之间有显著差异的话，那么这种将整个个案删除的方式将对调查结果产生严重偏差。因此采用这种处理缺失数据的方式时，一定要慎重。

① 柯惠新、丁立宏《市场调查与分析》，中国统计出版社，2000年，第198页。

3. 保留有缺失值的个案，仅在不同的分析中做必要排除（Paire-wise）

这种方法是保留有缺失数据的个案，多数情况下，将缺失值设立特殊的代码加以标记，比如，用“9”、“99”或“0”这样的数值来代表。标示后的缺失值在有些分析中要做必要的排除，比如，用“99999”来代表收入的缺失值，而要计算收入平均值的话，就要将“99999”这个值排除在外。这样做的结果，是在分析中不同的计算可能会基于不同的样本量来进行，那么在新闻报道中描述调查结果时，就要考虑到这个问题，在需要的时候，对受众做必要的交代。

在实践中，这种方法是应用得最多的，但要注意在应用中最好满足一定的条件。首先样本量要很大；其次缺失数据不多；第三变量之间不能是显著相关的。

在以上3种处理缺失数据的方法中，第三种最为常用。当然采用不同的缺失值处理方法，最后可能产生不同的结果。因此这个问题一定要仔细考虑、加以选择。

二、异常值的处理

在数据审核过程中发现的异常值，要追究它们的来源。如果来源于数据录入错误，那么在审核阶段就可以得到更正。如果不是录入错误，在对被调查者进行回访不太现实的情况下，可以考虑采用以下的方法加以处理。

1. 改变异常值的数值

改变异常值数值方法的思路是用某个估计量来代替异常值。有兴趣的读者可以参考其他书籍。

2. 调整异常值的权重

这种方法主要是降低异常值的权重，从而降低它们对结果



的影响。

三、对数据的加权处理

对数据进行加权处理，可能的目的有两个：其一是为了使样本更具有代表性；其二是为了强调某些被调查群体的重要性。

所谓加权处理，就是给数据集中的每一个个案或数据记录分配一个权重，用来表示这些个案或数据相对于其他个案或数据的重要性。加权处理反映在数据上的最终效果是增加（权重 >1)或减少（权重 <1）具有某些特征的群体的样本数量。

加权处理最常用的目的是使样本对于所研究的目标总体更具有代表性。例如，某报希望了解北京市男性居民在娱乐方面的花费情况，每个月的平均消费是怎样的。调查中没有采用随机抽样的方式，在回收的数据中，18～24 岁的男青年占 49%（见表 9-5)，因而样本是有偏的。但是研究者仍然希望能从中获得尽可能多的有关北京市男性居民每个月在娱乐上所花费用的信息。为了更好地做到这一点，在数据处理之前，根据北京市 18 岁以上男性居民总体的年龄分布，对样本进行了加权处理，结果如表 9-5。

表 9-5 通过加权增加样本的代表性

年龄	样本（%）	总体（%）	权重（总体比例/样本比例）
18～24 岁	49.0%		
25～34 岁	12.0%		
35～44 岁	28.0%		
45～54 岁	2.0%		
55 岁或以上	9.0%		
合计			

加权处理的另一个目的就是调整样本，为重要性大的被调查者赋予更大的权重，增加其相对重要性。例如，在研究有关城市治理的某项措施的意见时，研究者可能会认为专家的意见比一般群众的意见要重要一些，因此在进行数据处理之前，就要对专家这个被调查群体给予更大的权重。

在对数据进行加权处理一定要慎重，做之前要仔细考虑。如果真的需要做加权处理，那么在最终描述调查结果时，要清楚地加以说明。

四、原始数据或变量的转换

在对数据进行分析之前，可能需要对原始数据进行一定的处理，这里所说的处理包括两个方面内容：

1. 对原始数据重新分类、重新编码

很多时候，研究者在收集数据时，可能会采用比较容易的方式进行，比如，在询问被调查者的年龄、月收入时，问卷设计经常会采用定距变量的形式："请问您是在哪一年出生的？""请问您每个月的收入平均是多少？"等等。在数据采集回来之后，研究者可能会觉得将数据合并成新的类别，生成定类变量，这样分析起来更有意义。比如，可以将月收入分为5个类别：低收入、中低收入、中等收入、中高收入、高收入，等等。这时就需要将原始数据进行重新分类。

有的时候，我们在编码的时候，为避免遗漏某些重要的类别信息，会有意将编码表做得很细，类别很多，那么在数据处理之前，可能需要将数据合并成更少的几个大类别。这个时候也需要将原始数据重新分类，重新编码。

在做重新分类和重新编码的工作时，需要注意所有的被调查者数据都要包含在新的类别之中，并且新的类别没有交叉或



重叠，也就是说，新的类别应该是穷尽的和互斥的。

2. 重新定义变量或修改现有变量

重新定义变量或修改现有变量经常也是需要的。比如，某项调查想了解被调查者家庭的年收入，而访问时，为了减轻被调查者的回答负担，问题设计成："您家庭每个月的平均收入（包括工资、奖金及其他一切收入）是多少?"数据收集上来后，为了达到最初的研究目的，就要对家庭月收入这个变量（Income）进行计算，重新定义一个新的变量：T_income = Income * 12。新的变量 T_income 代表了被调查者家庭的年收入。

比较常见的定义新变量的情况，是为了使用某些分析方法，需要将定类变量转换成哑变量（0—1 变量），这样就可以近似按照数值型变量来处理。哑变量只有两个取值，一般规定为 0 或 1。如果一个定类变量有 K 个类别，那么就需要用 K—1 个哑变量来表示这个定类变量。例如，性别这个变量有两个类别，所以只需要用一个哑变量 X 来表示：

X = 1　　男性

X = 0　　女性

而教育程度有 3 个类别的话，就需要两个哑变量 X1 和 X2 来表示：

X1 = 1　　初中或以下　　X2 = 1　高中、中专

X1 = 0　　其他　　　　　X2 = 0　其他

为了更好地进行定量分析，很多定类变量都需要先转换成哑变量，如性别、学历、职业等等。

另一种常见的变量变换是标准化，对变量进行标准化变换的目的是使不同单位或不同量表测量出来的变量在分析中具有可比性。例如，使用某个量表测量的得分为 X，对 X 进行标准

化，就是用标准得分Z来代替X。其中X=〔X－E（X)〕/S。E（X）是变量X的均值，S表示X的标准差。利用各自的标准Z得分，可以比较不同类型量表测量的变量。

五、编制变量标签和变量值

在进行数据录入时，为了提高录入速度，一般都是以字母代表变量名称，变量的取值则直接录入数字。同样以表9-1为例，录入完成时得到的数据集，一般使由表9-1的第二列变量名称和表9-1的最后一列变量编码说明中的数字组成。在对该数据集中的数据进行统计分析之前，要对其中的变量和由这些变量定义出来的新变量的名称和取值进行详细说明，对名称的说明称为变量标签（Variable Label），简称变标，如表9-1的第三列；对变量取值含义的说明称为变量值标签（Value Label），简称值标，如表9-1的最后一列。这些值标和变标一般要做成可执行文件，由统计软件自动运行、读取。在完成变标和值标的编制任务后，要给出类似表9-1的变量说明，尤其对于一些新定义或新生成的变量，要详细加以说明。

在上述步骤完成之后，即得到了一个可以用于进行数据分析的数据集．有的时候，需要确定数据储存的格式。储存数据有两种主要格式：数据库和二进制数据文件。二进制数据文件ASCII码能用不同的制表软件或统计软件读取。但是使用ASCII码文件，大多数的统计软件还要重新生成自己特定的格式。如果数据以一种适当的数据库格式存贮，大多数统计和数据库软件就可以不需要重新生成文件而直接处理，查询操作也可以在数据库上直接进行。较好的选择是对调查结果生成一个二进制数据文件，同时生成数据库文件。



第十章　SPSS for Windows 基本操作

电子计算机的普及和发展，使得我们可以借助计算机快速地对数据进行处理，目前国际上有一些通用的统计软件包，如 SPSS 和 SAS 等，在这里，我们对统计软件 SPSS 做一简要介绍。使用 SPSS，可以有效地完成前面我们介绍的各种统计方法。

第一节　SPSS 的基本特点和功能介绍

随着 SPSS 软件版本的不断更新，软件的功能也不断完善，操作越来越简便，与其他软件的接口也越来越多。使用 SPSS 软件，可以在实现统计功能的同时，将分析结果用清晰简练的表格或（和）生动形象的二维、三维图形来呈现。

一、SPSS for Windows 的基本特点

与 SPSS/PC + 相比，SPSS for Windows 具有如下一些明显特点：

(1) 使用该软件，只要给出分析指令，系统便自动进行数据处理，得到相应的结果，为我们节省了做大量数学运算的时

间，我们所要做的，就是运用统计理论和方法来解释和阐述软件计算出的结果。

(2) SPSS for Windows 具有很多个统计功能模块，使其分析数据的功能非常强大。

(3) 完全的 Windows 风格可视化界面，输入数据文件以后，只需用鼠标结合简单的数据输入便可完成操作。

(4) 完善的帮助系统（包括图解帮助、在线帮助和联机帮助等），可以帮助自学该软件。

(5) 单击对话框中的“Paste”按钮，可以自动生成“Syntax”程序代码，经过简单的编辑或直接运行，就可以实现指定的功能。

(6) 该软件不仅仅有数理统计的基本功能，还具有丰富的图表表达功能，可以根据需要选择不同风格、不同功能的表格和图形来整理数据，使我们的报告、文章等更为美观。

(7) 可以与很多其他软件进行数据传输。SPSS 软件可以打开扩展名为 .dat、.xls、.slk、.dbf、.wk3 等多种数据文件。该软件生成的图形可以保存为多种图形格式。

二、Windows 环境下 SPSS 软件的使用

SPSS6.0 以上的版本均支持 Windows 环境下的操作，在 Windows 界面下，单击“开始”按钮，然后在打开的菜单中用鼠标指向“程序”选项，在打开的子菜单中单击“SPSS x.0 for Windows”选项，即可打开 SPSS 的对话框，进入 SPSS for Windows 的主环境。在此，我们将以 SPSS10.0 为例，介绍 SPSS 的基本操作。

进入 SPSS 主环境时显示的对话框提醒你下一步准备做什么，进行选择后，可以帮助你快捷地完成任务。如图 10-1 所示。

该对话框中包括以下选项：

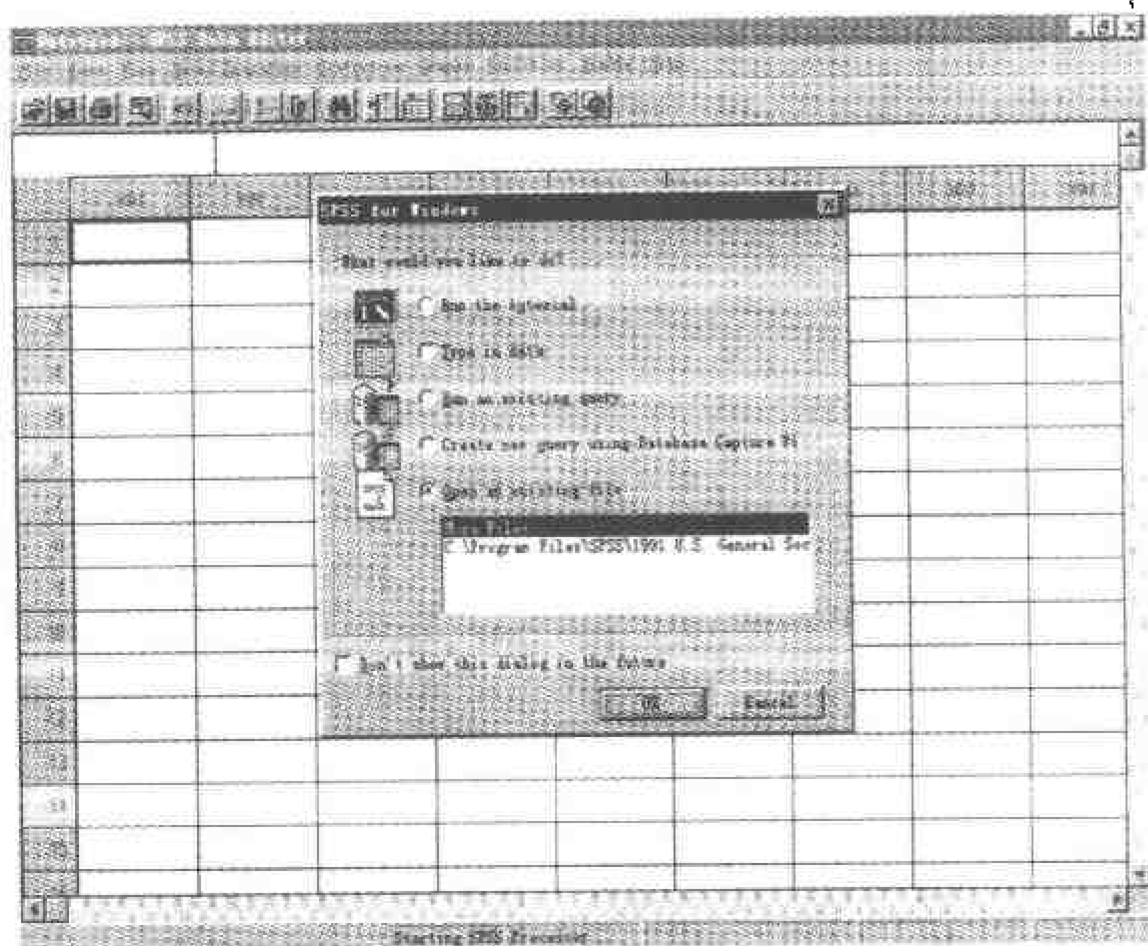


图 10-1　SPSS for Windows 对话框

1. Run the tutorial 选项，选择该项，将直接运行图解帮助。

2. Type in data 选项，选择该项，可在编辑器中直接输入数据。

3. Run an existing query 选项，选择该项，可以运行一个已经存在的数据库文件。

4. Create new query using Database Capture Wizard 选项，选择该项，用 Database Capture Wizard 应用程序创建新的数据库文件。

5. Open an existing file 选项，为缺省选项。选择该项，在下面的列表框中进行选择，打开一个已经存在的 SPSS 数据文件。在此列表框中双击“More Files…”选项，然后在打开的对话框中进行选择，可以打开更多的数据文件。

6. Don't show this dialog in the future 选项、选择该项，以后将不显示该窗口。

三、系统的运行、管理

1. SPSS 系统的运行方式

(1) 完全窗口菜单运行管理方式

SPSS for Windows 启动后在屏幕上显示主画面，即为数据编辑窗口，窗口的最上面是由 10 个菜单项组成的主菜单，如表 10-1 所示。

表 10-1 SPSS 主菜单项说明

	菜单名称	实现功能
1	File	文件操作
2	Edit	文件编辑
3	View	视图编辑
4	Data	数据文件建立与编辑
5	Transform	数据转换
6	Analyze	统计分析
7	Graphs	统计图表的建立与编辑
8	Utilities	实用程序
9	Window	窗口控制
10	Help	帮助



当然每个菜单项还包括一系列的功能。当在数据编辑窗口中输入或读入数据文件后，可以使用主菜单项的各种功能进行工作。单击选中的菜单项就可以打开菜单，然后或者进一步选择细分功能，或展开有关功能的对话框，通过对话框选择参与分析的变量，利用其中的功能钮进一步选择分析参数，以便执行选中功能。

（2）程序运行管理方式

程序运行管理方式是在语法窗口（Syntax）中直接运行编写好的程序的方式。在 Syntax 窗口中的程序可以是分析数据窗口中的数据，也可以用有关的语句指定外部数据文件进行分析。

（3）混合运行管理方式

混合运行管理方式是上述两种方法的结合。首先在数据窗口中输入数据或打开已经存在的数据文件，然后利用对话框选择分析功能、分析变量和分析参数，然后用“Paste”按钮将选择转换成相应的命令语句，粘贴在“Syntax”窗口中，在其中修改或添加语句、参数等，然后通过“Run”按钮，执行程序。

2. SPSS 中的窗口及其功能

SPSS10.0 中的窗口包括数据编辑窗口、输出窗口、语法窗口、统计图形编辑窗口、帮助窗口等。

（1）数据编辑窗口（Data Editor）

SPSS 主画面即为数据编辑窗口，可在该窗口中编辑数据文件。SPSS10.0 中数据编辑窗口分为两个视区：“Data View”和“Variable View”。图 10-2 为其中的“Data View”视区。该视区中有一个可扩展的平面二维表格。“Variable View”视区行为变量，列为变量名称、变量类型等内容（详见第十章第二节）。

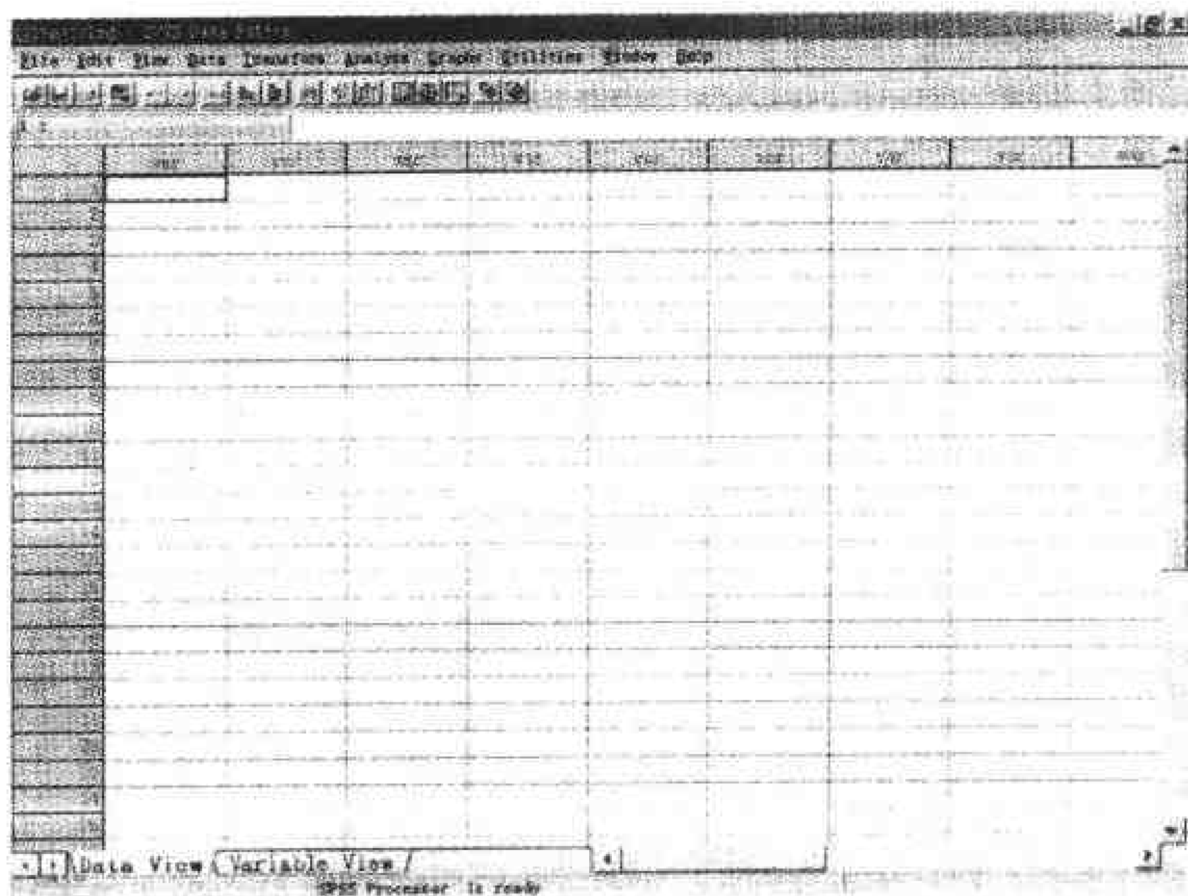


图 10-2 数据编辑窗口中的“Data View”视区

利用这个主画面下的系统菜单有“File”、“Edit”、“View”、“Data”、“Transform”、“Analyze”、“Graphs”、“Utilities”、“Windows”、“Help”。

其中“File”可以实现对文件的操作、打开、保存等。在SPSS运行期间，不能同时打开两个数据文件。

- “Data”项可以完成对变量属性的定义和修改，数据录入，移动记录指针，插入记录或变量，还可以对数据记录进行排序、转置等。
- “Edit”菜单项可以对数据进行复制、剪切、粘贴等。
- “View”菜单项可以进行视图编辑和窗口外观控制。
- “Transform”菜单项可以实现数据和变量变换、计算等。



• "Analyze" 菜单项可以对该窗口中的数据进行统计分析。

• "Graphs" 菜单项可以进行各种统计图表的建立和编辑。

• "Utilities" 菜单项含有变量列表、文件信息、定义与使用变量集合等。

• "Windows" 菜单项可实现所有窗口的最小化和最大化，激活窗口列表。

• "Help" 菜单项提供有关 SPSS 的各方面帮助。

SPSS 的数据编辑窗口不能关闭，关闭这个窗口就意味着结束 SPSS 的运行。

(2) 输出窗口 (Output - SPSS Viewer)

SPSS 的输出窗口如图 10-3 所示。当使用 "Analyze" 进行统计分析产生数据输出时，就可以激活输出窗口，在该窗口中显示 SPSS 处理的输出信息；另外通过 "File" 菜单项也可以打开一个新的输出窗口。

①输出窗口中的主菜单

输出窗口中的主菜单包括"File"、"Edit"、"View"、"Insert"、"Format"、"Analyze"、"Graphs"、"Utilities"、"Windows"、"Help"。

其中"Analyze"、"Graphs"、"Utilities"、"Windows"、"Help"中的功能和命令项与数据编辑窗口中的相同。

• "File" 菜单项中增加了关闭窗口、保存时设置密码、输出、页面设置、打印预览等功能。

• "Edit" 菜单项中增加了选择、全选、特殊粘贴等功能。

• "View" 菜单项包括显示/隐藏切换、字体设置等功能。

• "Insert" 菜单项包括插入/删除分页符、图表编辑功能等项。

• "Format" 菜单项包括对齐功能、行列装置、页、行、列转置功能和调整表格各元素尺寸等功能。

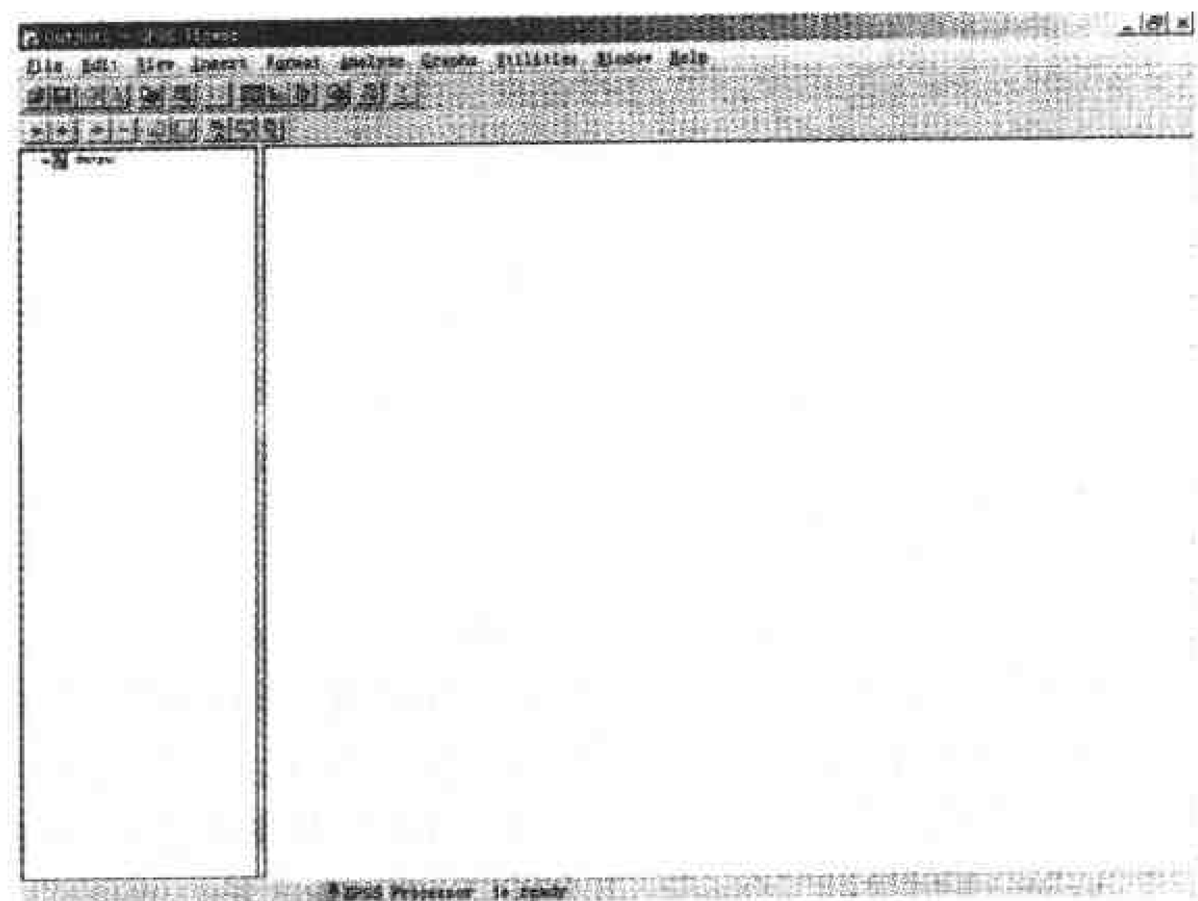


图 10-3　SPSS 输出窗口

②功能图标按钮

功能图标按钮包括输出信息操作功能图标按钮，如，打开文件、保存文件、打印及预览、输出、查找数据、显示变量信息等；还包括导航图标功能按钮，如，向前定位输出组，向后定位输出组、插入表头、标题、文本等按钮。图标按钮还可以根据需要增减。

③输出文本窗口

输出窗口的右半边是一个文本窗口，可对其中的内容进行编辑。

④输出导航窗口

输出窗口的左半边是一个导航窗口，可对其中的内容进行

编辑。

⑤关闭输出窗口

SPSS允许同时建立多个输出窗口，而主输出窗口只有一个，各过程输出结果会输出到主窗口。当输出窗口不是主输出窗口，并且是当前工作的输出窗口时，可以关闭该输出窗口。

(3) 语法窗口 (Syntax)

通过主菜单中的“File”菜单项，可以打开Syntax窗口，如图10-4所示。



图10-4　Syntax窗口

“Paste”按钮在几乎所有的SPSS对话框中都存在，它是专门为编程准备的，作用是把在对话框中所做的SPSS过程的命令语句，以及各选择项对应的子命令语句按照SPSS语言的语法组成一个或若干个完整程序粘贴到Syntax窗口上。

除了使用“Paste”按钮将过程语句粘贴到 Syntax 窗口之外，还可以使用键盘输入 SPSS 命令编写的程序，并进行编辑、修改、保存等。SPSS 语句一条可以占用多行，最前面为语句主体，具体的选项用斜杠和语句主体相连，最后用小数点结束语句。当选定一个完整的程序后，可以用“Run”按钮，提交系统执行选定程序。鼠标对准“Syntax Help”按钮单击，屏幕上会显示光标所在行上的 SPSS 命令语句供查阅。

在 SPSS 中也可以同时打开多个语法窗口，可以关闭激活的当前工作语法窗口。

(4) 图形编辑窗口 (Charts Editor)

Chart Editor 窗口是一个图形编辑窗口，在该窗口中可以对图形进行编辑、修改、打印。该窗口如图 10-5 所示。

(5) 帮助窗口

几乎每个对话框或大部分子对话框中都有一个“Help”按钮，单击就可以打开一个帮助窗口。窗中显示的帮助信息与当时的操作有关，或与当时打开的对话框中的内容有关。

3. SPSS 中常见对话框类型

主菜单中各项功能的完成总要通过对话框选择命令、语句、变量或者参数，提交系统执行，才能得到要求的结果。通过各种选择按钮、下拉菜单等指定选择项、执行命令、打开子对话框。

SPSS 中使用的对话框主要有两种：文件操作对话框和统计分析主对话框。

(1) 文件操作对话框

可从主菜单“File”中打开的各种功能的对话框，例如，打开数据文件、保存文件、打印等功能项的对话框。



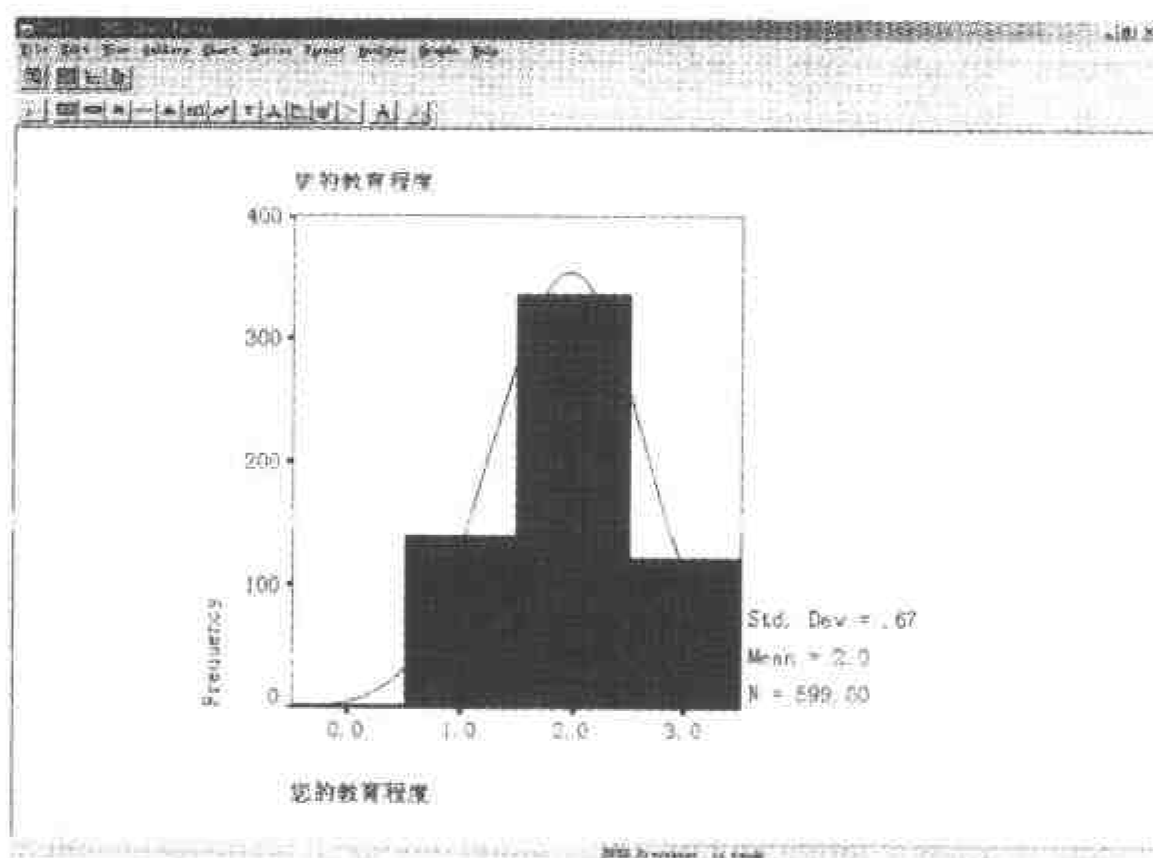


图 10-5　图形编辑窗口

（2）统计分析主对话框

主菜单“Analyze”的每个子菜单对应于一个过程分析。进入子菜单打开的第一个对话框均为统计分析对话框，该对话框主要用来选择参与分析的各类变量，另外还有一些其他可选项，用来选择算法，输出选择项等。

（3）其他选择项对话框

其他选择项对话框包括主菜单的其他菜单项各功能对应的对话框或统计分析菜单项各分析过程的二级对话框。

四、SPSS for Windows 的功能介绍

SPSS for Windows 可以实现数据编辑、表格生成；编辑、

图形生成；与其他软件的联接以及各种统计功能。

1．SPSS的数据编辑功能

在SPSS的数据编辑窗口中，不仅可以对打开的数据文件进行增加、删除、复制、剪切和粘贴等常规操作，还可以对数据进行排序、转置、拆分、聚合、加权等操作，对多个数据文件可以根据变量或个案进行合并。可以根据需要把待分析的变量定义为一个集合，打开时指定打开该集合。

2．表格的生成和编辑

利用SPSS可以生成数十种风格的表格，根据表格的功能又可分为一般表、多重响应表和频数表等。利用专门的编辑窗口或直接在输出窗口中可以编辑生成的表格。

3．图形的生成和编辑

利用SPSS，可以生成数十种基本图和交互图。其中基本图包括条形图、线图、面积图、饼图、帕累托图、箱图、散点图、直方图、P-P概率图、Q-Q概率图等。交互图包括条形交互图、点形交互图、线形交互图、直方交互图、散点交互图等。图形生成以后，可以进行编辑。

4．与其他软件的联接

SPSS能打开Excel、DeBase、FoxBase、Lotus 1-2-3、Access、文本编辑器等生成的数据文件。SPSS生成的图形可以保存为多种图形格式。

5．SPSS的统计功能

SPSS的基本统计功能包括：

(1) 对样本数据的描述和预处理；(2) 假设检验（包括参数检验、非参数检验及其他检验)；(3) 方差分析；(4) 列联表（交互表)；(5) 相关分析；(6) 回归分析；(7) 对数线性分析；(8) 聚类分析；(9) 判别分析；(10) 因子分析；(11)



对应分析；（12）时间序列分析；（13）生存分析；（14）可靠性分析。

五、SPSS for Windows 的帮助系统

SPSS 提供了多种帮助手段，利用这些帮助，可以很方便地实现查询和解疑。

1．主菜单“Help”提供的帮助功能

（1）Topics

帮助主题：SPSS for Windows 对话框如图 10-6 所示，它是关于各种主题的帮助，可以用目录查询，也可以输入关键字检索。打开某一主题后的帮助对话框如图 10-7 所示。

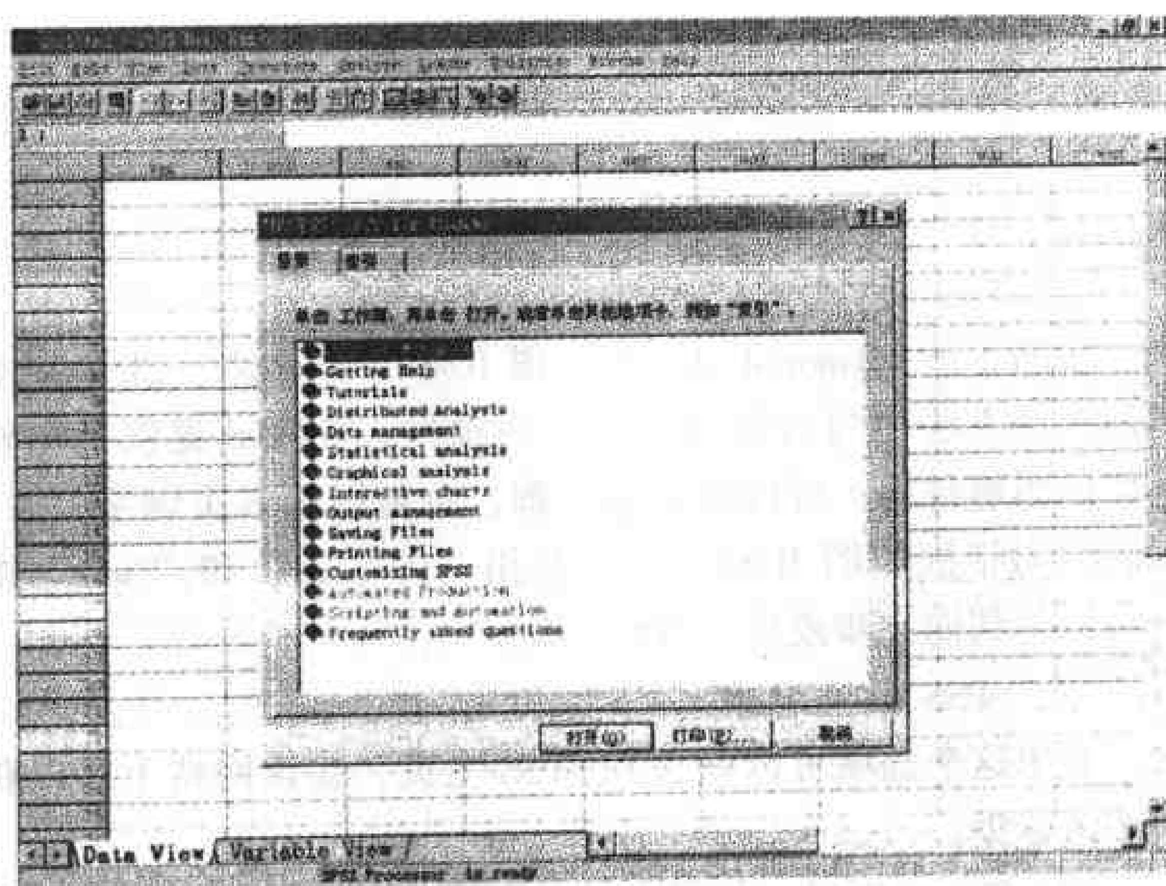


图 10-6　帮助主题：SPSS for windows 对话框

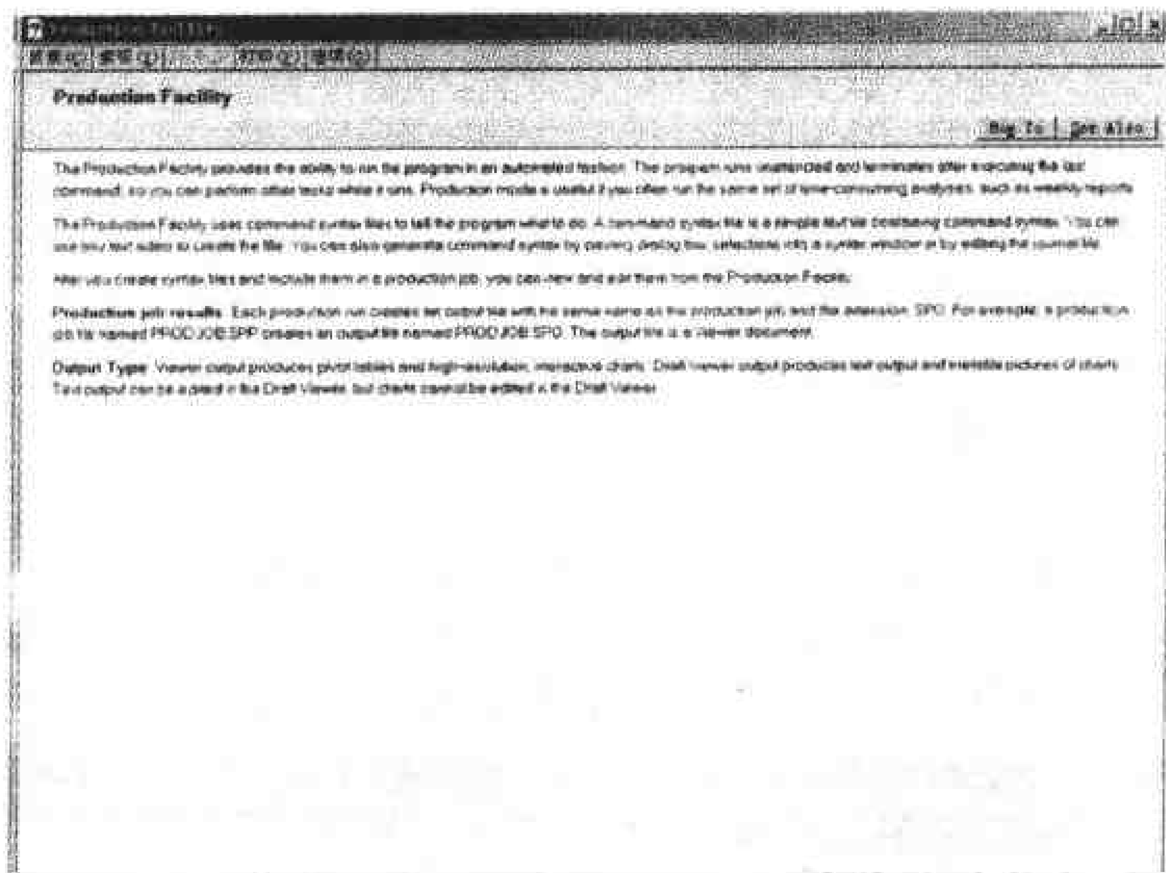


图 10-7 SPSS for windows 对话框

（2）Tutorial

帮助主题：Tutorial 对话框如图 10-8 所示。对于 SPSS 中基本的操作方法和与数据有关的一些操作，Tutorial 提供了比较具体的图解帮助。可以用目录查询，也可以输入关键字检索。Tutorial 对话框如图 10-9 所示。利用“Previous”和“Next”按钮可以得到前一步或后一步的帮助。

（3）SPSS Home Page

利用这个选项可以联接到 SPSS 主页，在该网站上可以获得在线帮助。

（4）Syntax Guide（语法指南）

语法指南的下一级菜单有 3 项：Base（基础部分）、Regression Model（回归模型）、Advanced Model（高级模型）。



图 10-8　帮助主题：Tutorial 对话框

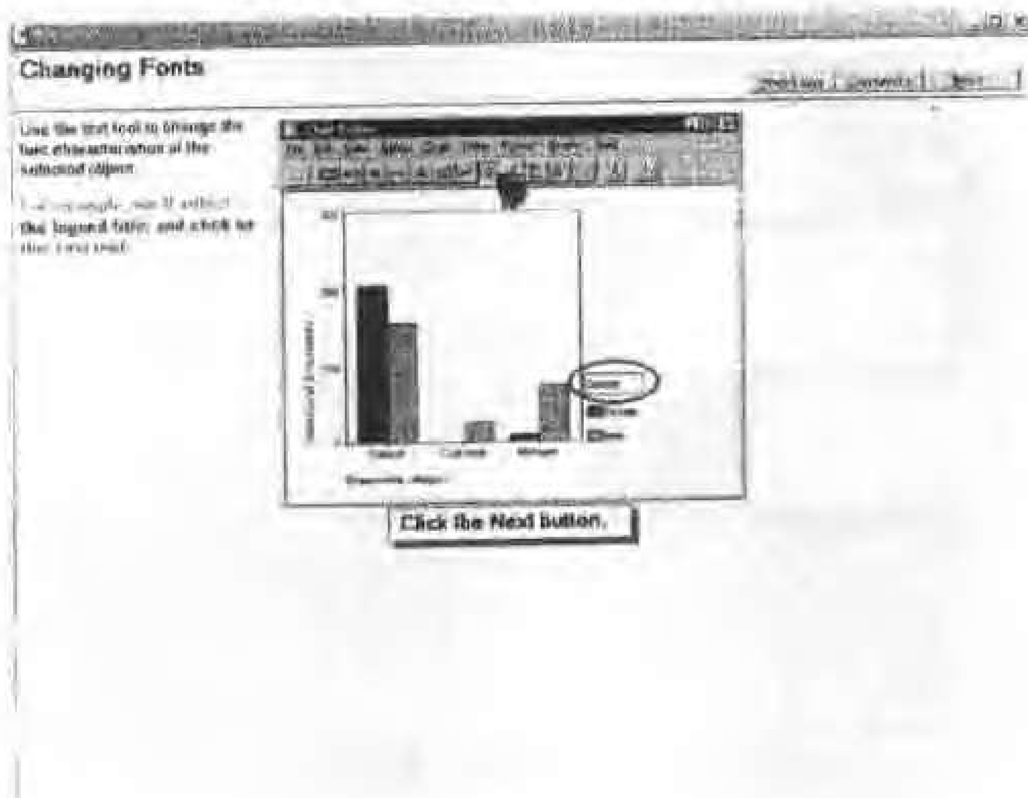


图 10-9　Tutorial 对话框

（5）Statistics Coach（统计分析指南）

统计分析指南对话框如图 10-10 所示。从右侧选择所要进行的分析，左侧即有相应的帮助和指导，并可得到更多的范例。

2．在线帮助

在 SPSS 中操作，经常要打开一系列对话框，这些对话框中一般有一个“帮助”（或“Help”）按钮，可以提供在线帮助。单击该按钮，可以打开与本对话框内容有关的帮助信息。例如，你想利用数据做 High-Low 图，要获得帮助，只需单击 High-Low 对话框中的“Help”按钮，即可打开相关的帮助信息，从中可以了解创建方法、对话框中各控件的意义等。如图 10-11 所示。

3．右键提示

在使用 SPSS 的过程中，当对对话框中某个控件的用途不清楚时，单击鼠标右键，便可以获得相应的帮助信息。图 10-12 为 Linear Regression 对话框中“Independent”列表框的右键帮助信息。

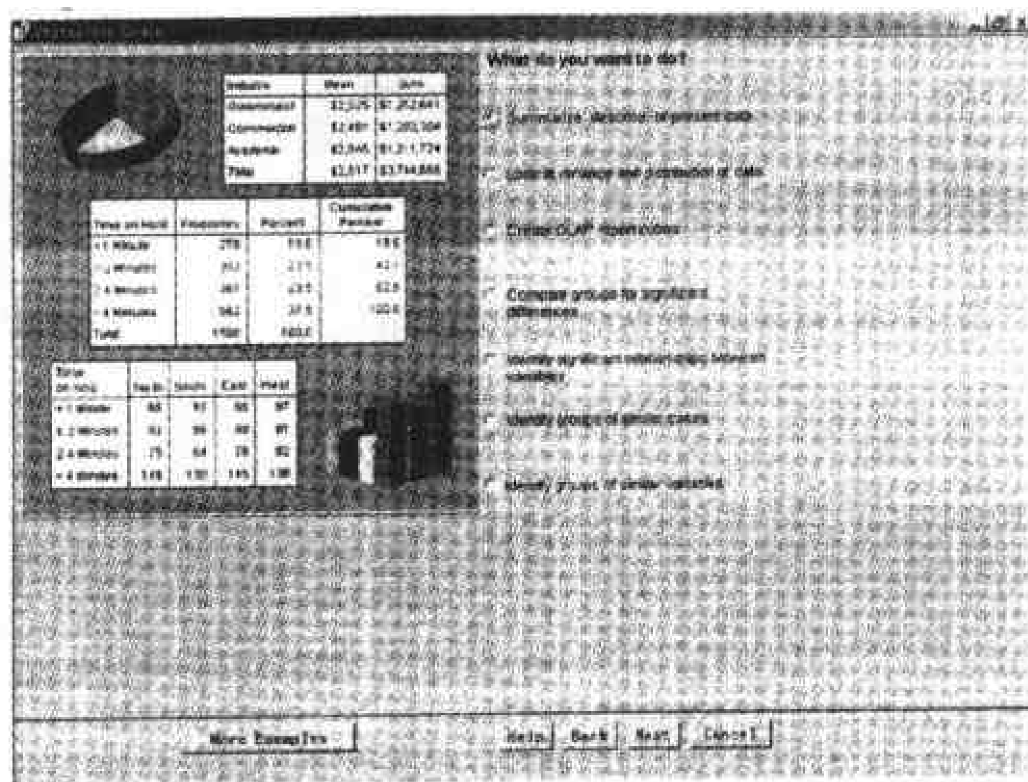


图 10-10　Statistics Coach 对话框

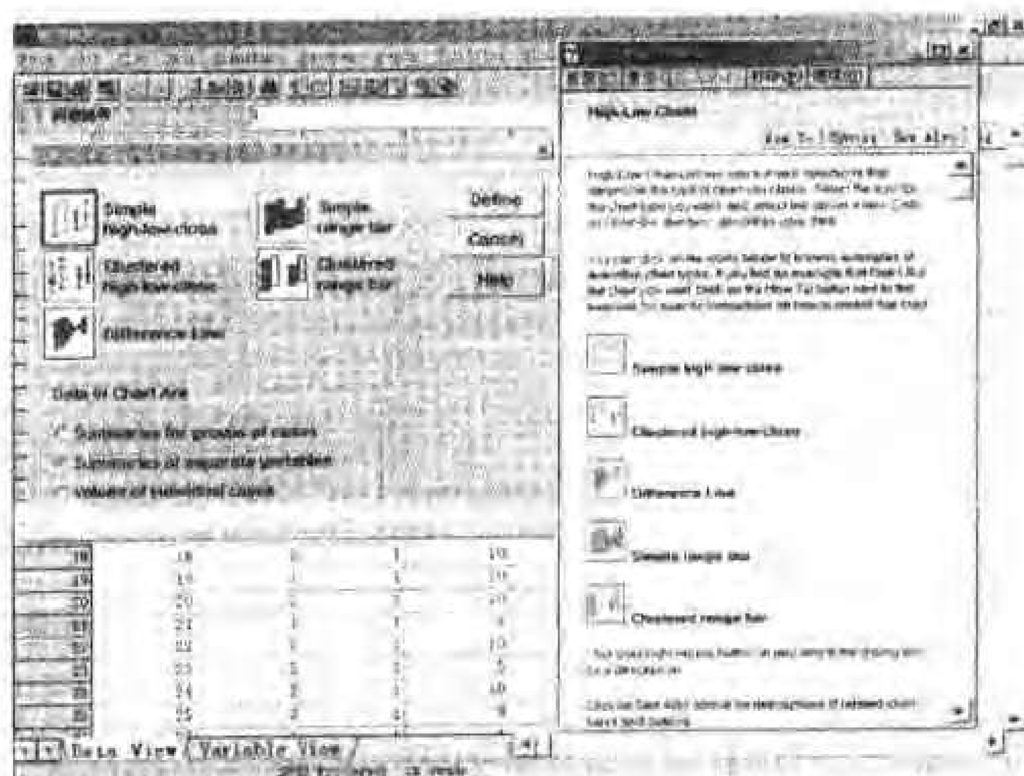


图 10-11　High-Low Chart 对话框及其帮助信息

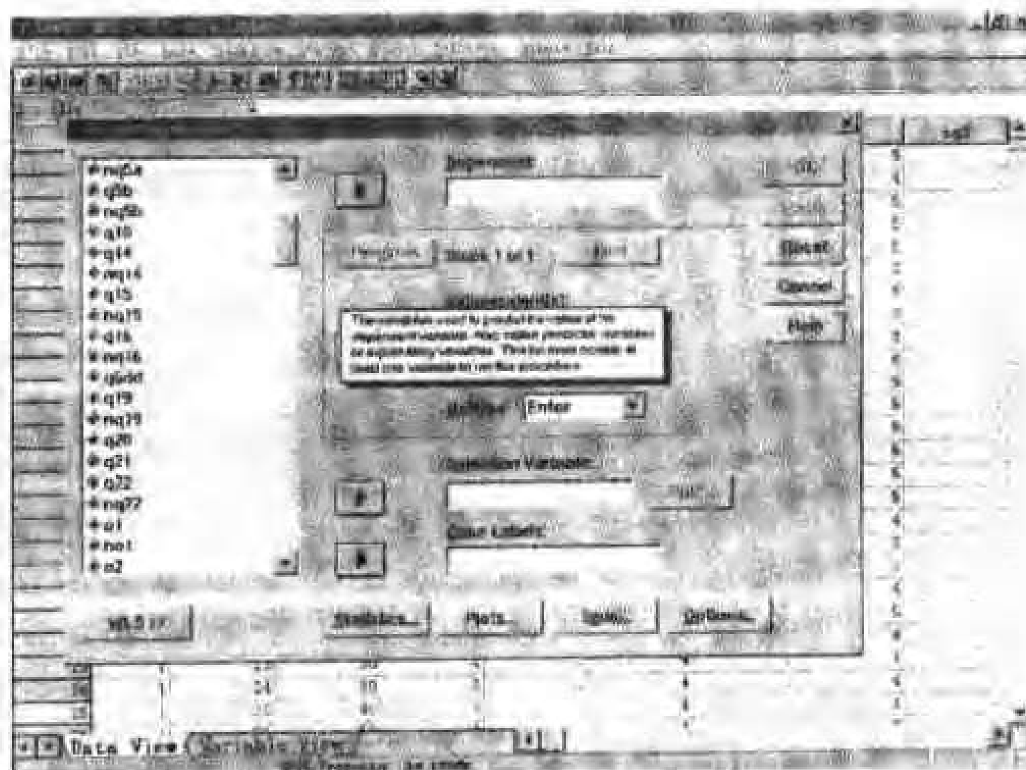


图 10-12　右键帮助信息

第二节 变量及其相关操作

SPSS 中的变量有 4 个属性，即变量名称（Name）、变量类型（Type）、变量标签（Label）、变量长度（Width）。在定义 SPSS 变量时，至少要定义变量名称和变量类型，另两个属性可以使用 SPSS 默认值。

一、关于变量

1. 变量名称命名规则

SPSS 变量名称的命名规则包括以下几点：

（1）变量名称的第一个字符必须为字母，后面可跟任意字母、数字，或除“?”、“!”、“*”、“ ”和空格以外的字符；

（2）变量名称不能以下划线“-”或句点“.”结尾；

（3）变量名称的长度一般不能超过 8 个字符；

（4）SPSS 系统不区分变量名称中的大小写；

（5）下面的关键词是 SPSS 的保留字，不能用做变量名称：

ALL　NE　EQ　TO　LE　LT　GE

BY　OR　GT　AND　NOT　WITH

按照以上规则，ABCD _ 1234、ABCD.、1ABCD、ABCD?1、ALL 等都是不合法的变量名称，而对于 SPSS 来说，ABCD 和 abcd 是同一个变量。当在变量定义窗口中输入不合法的变量名称时，系统会自动显示出错信息。

2. 变量类型和长度

SPSS 变量有 3 种基本类型：数值型、字符型（String）、日期型（Date）。而其中数值型变量又分为 5 种，即标准数值型变量（Numeric）、逗号数值型变量（Comma）、句点数值型变量



(Dot)、科学记数型变量（Scientific notation)、带美元符号的数值型变量（Dollar)。

在对精确新闻报道题材的分析中，最为常用的是标准数值型变量。这是最常用的一种变量类型，也是 SPSS 中变量的缺省类型。这类变量在 SPSS 中的默认总长度是 8 位，其中小数位数占两位。标准数值型变量的值用标准数值格式显示，可以在定义的时候指定其长度，否则就使用默认长度。当定义的小数位数小于原数据的小数位数时，将按照四舍五入的原则收尾。定义为标准数值格式的变量值可以用标准数值格式输入，也可以用科学记数法输入，但此时显示出来的格式还是标准数值格式的。

3. 变量标签与变量值标签

我们知道，变量与变量的值是两个不同的概念。变量用来描述各个个案（样本）的特征和属性，对于不同的个案而言，变量的取值会有所不同，是为变量；而变量的值是某个特定的个案的取值，对这个个案而言，它是固定的，是常量。

(1) 变量标签（Variable Label)

变量标签是对变量名称所做的进一步说明。在 SPSS 中，变量名称最多只允许 8 个字符，通常 8 个字符不能明确、完整地表示变量的含义，这就需要应用变量标签。当为变量定义了变量标签以后，SPSS 在对该变量进行数据处理后得到的图表就用变量标签来表示标量的信息，否则，将直接用变量名称来表示变量的信息。表 10-2（a）和表 10-2（b）是对某个数据文件中的 Q2d' 变量（用来描述被调查者的性别）进行频数分析得到的，其中表 10-2（a）没有加变量标签，而 10-2（b）则加了中文变量标签“性别”。在这两张表中，10-2（b）让人一看就明白什么意思，而 10-2（a）单从变量名称看，就不会明白

它所代表的意思。

表 10-2 SPSS 加变量标签前后的表格输出比较

Statistics

Q2DD

N	Valid	200
	Missing	0

(a)

Statistics

性别

N	Valid	200
	Missing	0

(b)

SPSS 允许变量标签字符数最多为 256 个。

(2) 变量值标签 (Value Label)

变量的值标签是对变量可能取值所做的进一步说明，说明变量的每个可能取值所代表的含义。在定义了变量的值标签之后，可以使 SPSS 的输出结果更清楚，更便于阅读和理解。还以表 10-2 中的变量 Q2DD 为例进行频数分析。得到的表 10-3 (a) 没有加变量值标签，而表 10-3 (b) 则是加了值标签 (1. "男性", 2. "女性") 之后得到的 SPSS 频数输出。很明显，加了变量值标签后的表格输出更加清楚，让人一目了然。

表 10-3 SPSS 加变量值标签前后的表格输出比较

性别

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	102	51.0	51.0	51.0
	2	98	49.0	49.0	100.0
	Total	200	100.0	100.0	

(a)



性别

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	男性	102	51.0	51.0	
	女性	98	49.0	49.0	51.0
	Total	200	100.0	100.0	100.0

(b)

二、与变量有关的操作

1. 定义变量

输入数据之前首先要定义变量。SPSS 对变量的定义提供了很多种方法，可以在变量窗口进行定义；也可以通过建立和应用模板来定义。这里主要介绍通过变量窗口定义变量的方法。

定义变量就是要定义变量名称、变量类型、变量长度、变量标签和值标签以及变量的格式（如显示宽度、对齐方式、缺失值标记等)。以 SPSS10.0 为例，定义变量时，光标对准主窗口 Data Editor 下左下方的“Variable View”按钮，单击左键，即可打开定义变量对话框，如图 10-13 所示。

(1) 定义变量名称

由图 10-13 显示，定义变量窗口第一列为 Name，即变量名称，下面是系统默认的第一个变量 var00001。若使用系统默认的变量名称，且不定义其他的属性，则按回车键确认，或按 OK 键确认。这时变量描述框中显示的是变量的默认属性：

- Type：Numeric　　变量类型：数值型
- Width：8　　变量长度：8
- Label：None　　变量标签：(无)

- Missing：None　　缺失值：无定义
- Columns：8　　变量所占列数：8
- Align：Right　　对齐方式：右对齐
- Measure：Scale　　测量：定比例
- Decimals：2　　小数位数：2
- Values：None　　变量值标签：（无）

返回主窗口，系统按照变量的定义顺序给出的变量名称默认值为 var00001、var00002、var00003……依此类推。

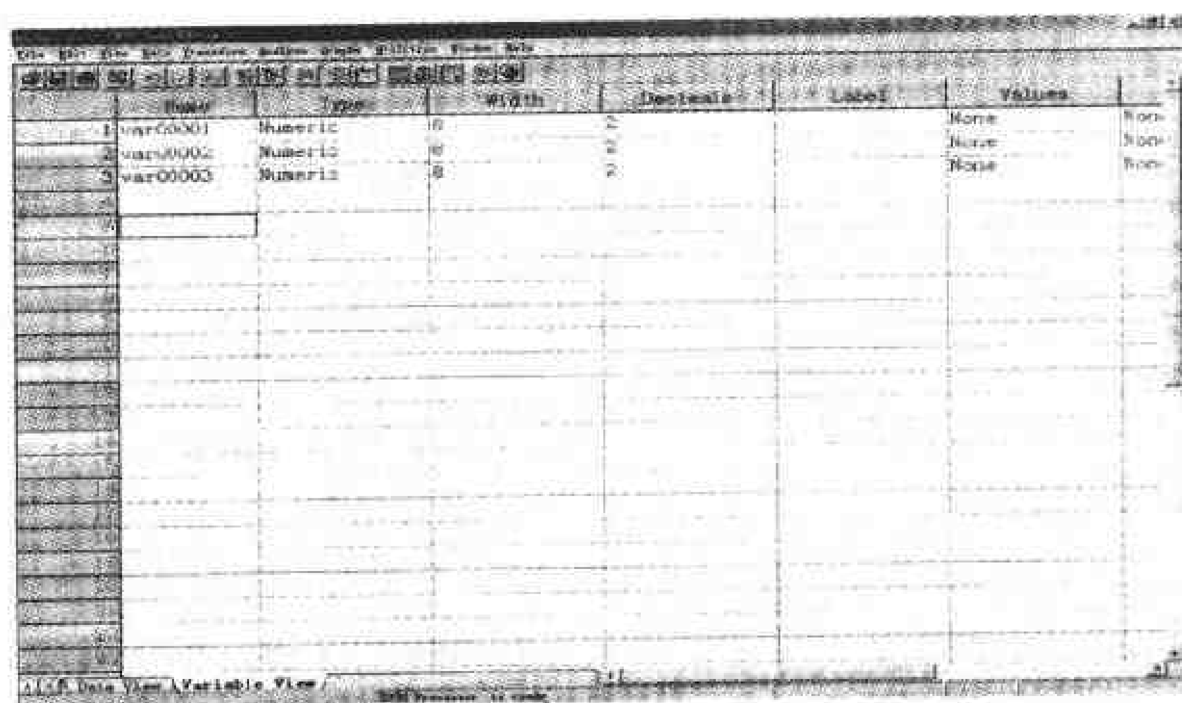


图 10-13　变量定义窗口

很多时候，操作者希望自己定义有一定意义的变量名称，那么只要将光标放在 Name 一列相应的单元格中，直接输入自己定义的变量名称即可。例如输入“Gender”作为变量名。

(2) 定义变量的类型与宽度

点击 Type 列，打开定义变量类型对话框，对话框中有 8 种可供选择的变量类型，其中 SPSS 系统默认的类型是标准数值型，即 Numeric 变量，也是最常用的变量类型。如果需要修



改，可直接选择其他类型。

定义变量的长度，在定义变量类型对话框中，显示当前选中变量类型的默认总长度和小数位数，如标准数值型变量，其默认总宽度是 8，小数位数是 2，如果要修改，可直接删除默认值，输入合适的值。例如，刚才定义的“Gender”变量，其取值只能是 1 或 2，因此其变量长度只需要一个字节，那么就可以把变量总长度改为 1，小数位数为 0。

(3) 定义变量标签和变量值标签

在定义变量窗口 Label 一列，对应要定义变量的单元格中，可直接输入定义的变量标签。例如，对前面定义的变量“Gender”，可输入“性别”作为它的变量标签。

点击 Value 列，可打开变量值标签定义对话框，对话框中第一个空格用来输入变量值，第二个空格输入对应的值标签，然后按 Add，即将该值标签加入到值标签清单中了。值标签定义后，可以修改，修改后只要按 Change 键就可以了。

(4) 定义变量格式

变量的格式包括变量的显示宽度、对齐方式、缺失值标记等。Column 列显示当前变量所占的列宽度。SPSS 系统默认值是 8，这个数值可以根据需要进行修改。

变量显示时的对齐方式有 3 种：Left（左对齐）、Right（右对齐）、Center（居中）。对数值型变量，系统默认值是右对齐，对字符型变量，系统默认值是左对齐。如果希望修改对齐方式，只需要用鼠标进行选择。

定义缺失值可以通过 Missing Value 对话框进行，系统默认的方式是无缺失值。如果想自己定义缺失值，可以采用几种方式：

第一种：离散缺失值（Discrete missing values），采用这种

方式，最多可定义3个离散的值为缺失值。如对前面定义的“性别”变量来说，取值只能是1或2，而3，4，5都是非法的，可以把这3个值定义为缺失值，在进行数据分析时，系统遇到这几个值的时候，就作为缺失值处理。

第二种：定义缺失值范围（Range of missing values），用这种方式定义缺失值主要针对连续变量的值，指定某一范围的值，如果出现在相应的变量取值中，则被当做缺失值处理。

第三种：一个范围加一个离散值（Range plus one missing value），例如，在“收入”变量取值中，50元、20001元、23000元、24500元都是错误数据，那么只要定义一个缺失值范围20001～24500，再加上一个离散缺失值50就可以了。

2. 变量的增加、转移、复制、删除和插入

(1) 增加变量

增加变量实际上就是定义新变量，操作方法如前所述。

(2) 变量的删除

在数据编辑器中删除变量可以采用两种方法：其一是使用鼠标选中该变量列，然后单击右键，选择Clear选项，即可删除选定变量列；其二可以利用Edit菜单下的Clear选项来删除变量。

(3) 变量插入

如果需要在现有变量中间插入一个新的变量，可以使用Data菜单下的Insert Variable选项，或用鼠标右键单击选定变量列，在弹出式菜单中选择Insert Variable选项，则所选变量列的左侧就会插入一个新变量列，SPSS赋予缺省变量名。如图10-14所示。

(4) 转移变量

转移变量实际上是一个剪切再粘贴的过程，即在选定变量

列后，先选择剪切，然后在目标列上粘贴。如果希望把选定列转移到已存在的两个变量列之间，必须先插入一个变量列，再把选定列的内容粘贴到插入的变量列中，否则会覆盖原有内容。

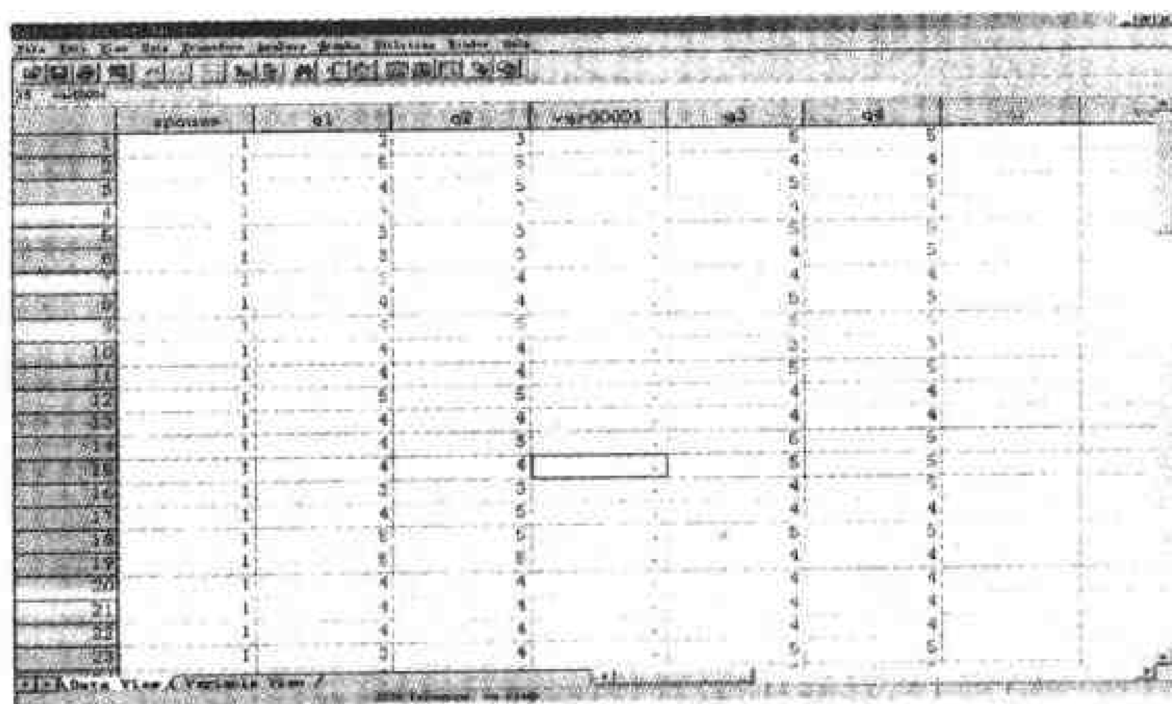


图 10-14 插入变量

(5) 复制变量

复制变量的方法与转移变量大致相同，不同在于复制会在原变量的位置上保留原变量，而转移则在原变量的位置上删除原变量。

第三节 数据输入和数据文件编辑

图 10-2 也是 SPSS 数据编辑器(Data Editor)的主窗口，其中，图中的数据列称为变量(Variables)，数据行称为个案(Cases)，各列上方是变量名称。

一、数据的输入

SPSS提供了多种数据输入方式，可以直接从数据编辑器输入，也可以从不同格式的ASCII数据文件中读入数据，或从Excel等其他数据库文件中读入数据。

1. 直接从数据编辑器输入数据

定义了变量之后就可以开始在数据编辑器中输入数据了。我们在图10-2中看到，数据编辑器是一个二维表格，顶部标有变量名称，左侧标有个案序号。一个变量名称和一个个案序号就定义了惟一的一个单元格。我们在这个单元格中就可以输入和这个序号个案中对应某个变量的取值。

SPSS允许多种输入操作方式，一般调查的数据是按照个案来输入的，即输入完一个个案在各个变量上的取值，再输入下一个个案。在一个个案数据的输入中，可以使用Tab键方便顺序激活一个个案中的单元格进行数据输入。

例：定义以下变量NO、X、Y、Z，下表是变量及其取值，将数据输入SPSS数据编辑器。

NO	X	Y	Z
1	1600	0.32	72
2	1450	5.40	67
3	900	3.42	76
4	1100	3.15	81
5	950	0.97	70
6	890	2.18	62

首先定义全部变量，如图10-15所示。

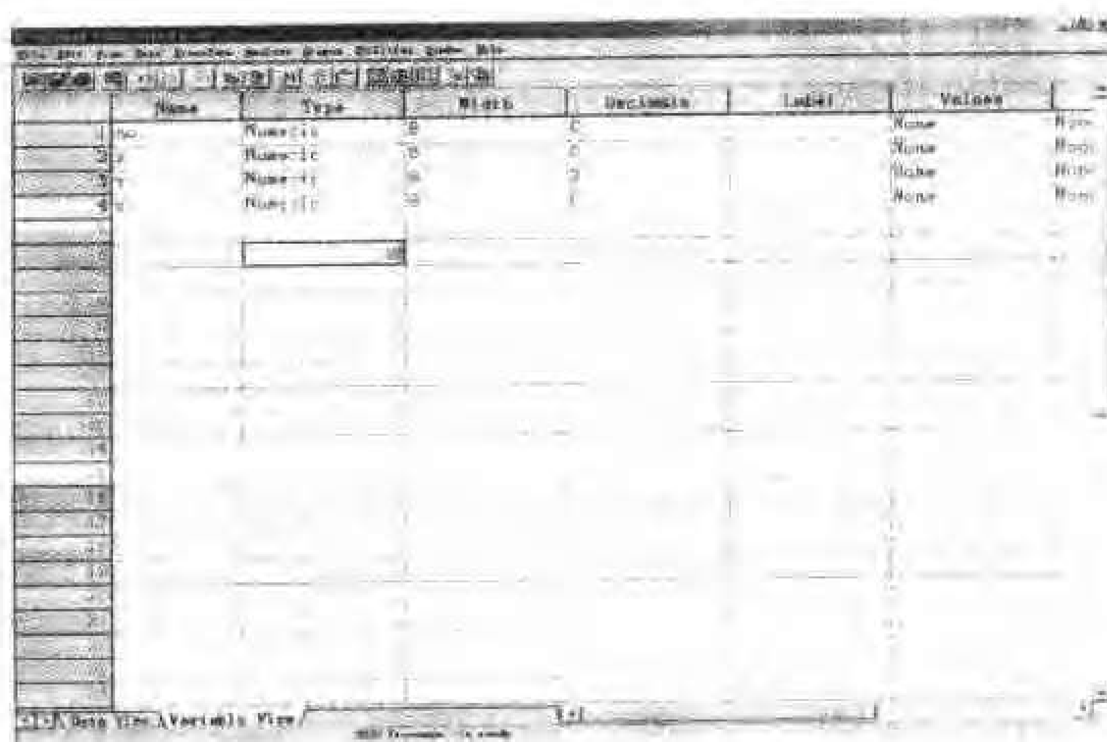


图 10-15　变量定义

然后在数据编辑窗口中输入数据，如图 10-16 所示。

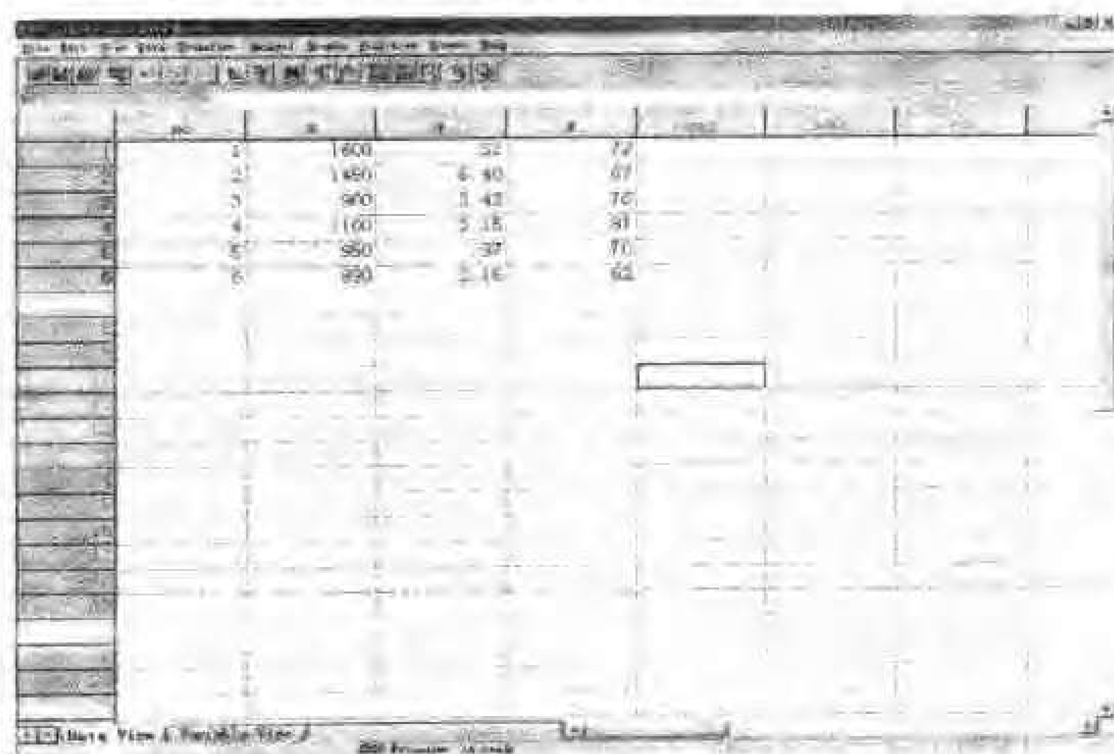


图 10-16　在数据编辑器中输入数据

2. 用 Open File 对话框输入数据

这种方法是通过 File 主菜单下的 Open data 选项，打开 Open File 对话框，利用这个对话框可以打开的文件格式包括：

- SPSS(* .sav)：是 SPSS for Windows 版本建立的数据文件
- SPSS/PC + (* .sys)：是 SPSS 的 DOS 版本建立的数据文件
- Systat (* .syd， * .sys)：一种数据文件
- SPSS Portable (* .por)：由 SPSS 建立的一种 ASCII 码数据文件，可以用于其他应用软件，其中变量类型、变量标签、值标签以及用户定义的缺失值都另外单独保存着。
- Excel (* .xls)：是 Excel 建立的表格数据文件
- Lotus (* .w *)：Lotus 1-2-3 格式的数据文件
- SYLK (* .slk)：表格数据文件
- dBase (* .dbf)：用各种版本的 dBase 和 FoxBase 创建的数据文件
- Tab-delimited (* .dat， * .txt)：ASCII 码数据文件

3. 输入 ASCII 码数据文件

ASCII 码数据文件是常用的一种数据文件格式，有自由格式和固定格式两种形式。自由格式对文件中数据间的间距、位置等没有固定的要求，输入数据时只要注意数据的先后顺序。固定格式要求各个变量所对应的数据占据固定范围的列，在此范围内数据可以任意放置。例如，某个变量的列位置是 5-10，而变量某一取值为 100，则 100 放在 5-8 列，或 7-9 列，或 8-10 列上都可以。

在 SPSS 中输入固定格式 ASCII 码数据文件数据文件的语法命令为：

DATA LIST FILE = ' 文件存储路径' FIXED/

在 SPSS 中输入自由格式 ASCII 码数据文件数据文件的语



法命令为：

DATA LIST FILE＝' 文件存储路径' FREE/

二、数据与统计结果的输出

1. 数据输出

在 SPSS 中对原始数据进行处理以后，可以以一定的格式输出。利用主菜单"File"中的"Save"或"Save as"选项保存数据文件。

SPSS 保存数据文件的对话框中可供选择的数据文件格式有以下几种：

• SPSS（*.sav)：保存为 SPSS for Windows 数据文件

• SPSS 7.0(*.sav)：保存为 SPSS for Windows 7.0 数据文件

• SPSS/PC+（*.sys)：保存为 SPSS DOS 版本数据文件

• SPSS Portable（*.por)：保存为 ASCII 码数据文件，可以用于其他应用软件，其中变量类型、变量标签、值标签以及用户定义的缺失值都另外单独保存着

• Table-delimited（*.dat)：保存为 ASCII 码数据文件，用回车换行符号作为两个观测量之间的分隔符

• Fixed ASCII（*.dat）保存为 ASCII 码数据文件

• Excel（*.xls)：保存为 Excel 软件可以读取的表格数据文件，每个观测量占一行，每个变量占若干列。用此种格式保存数据文件，变量标签、值标签以及缺失值定义均丢失

• 1-2-3 Rel3.0（*.wk3)：保存为 3.0 版本的 Lotus1-2-3 可读取的数据文件

• 1-2-3 Rel2.0（*.wk1)：保存为 2.0 版本的 Lotus1-2-3 可读取的数据文件

• 1-2-3 Rel1.0（*.wks)：保存为 1.0 版本的 Lotus1-2-3 可

读取的数据文件

- SYLK（*.slk）：保存为表格数据文件
- dBaseIV（*.dbf）：保存为 dBase 的数据库文件
- dBaseIII（*.dbf）：保存为 dBase 的数据库文件
- dBaseII（*.dbf）：保存为 dBase 的数据库文件

其中缺省的保存格式为 SPSS（*.sav）格式。

2. 统计结果的输出

（1）统计输出内容的保存

对数据进行统计运算，其结果显示在 SPSS 的输出窗口中，同样可以利用主菜单“File”中的“Save”或“Save as”选项保存输出内容，缺省的文件类型为 Viewer Files（*.sav）。

（2）输出结果的编辑

在 SPSS 输出窗口中，可以选定某个结果，双击进行编辑。

选定某一统计结果后利用主菜单中“Edit”或单击鼠标右键的“Copy”或“Copy Objects”选项，可以复制输出结果（表格或图形），然后在其他软件，如 Word 中粘贴。表 10-4 是从 SPSS 输出结果中粘贴过来的表格。

表 10-4　SPSS 输出的频数统计结果

性别

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	90	45.0	45.0	45.0
	2	98	49.0	49.0	94.0
	99	12	6.0	6.0	100.0
	Total	200	100.0	100.0	

（3）统计结果的打印



用户可以将输出文本打印，只需激活需要打印的窗口，选定需要打印的内容，使用和其他软件相同的打印命令，实现打印功能。

三、数据文件的编辑

在使用 SPSS 对数据进行统计分析之前，可能需要对数据进行整理、分类、变换或计算，也就是对数据文件进行编辑。很多时候，这是进行数据分析的基础。数据文件的编辑主要用到主菜单中的“Data”和“Transform”两个菜单项，包括的内容：数据的排序；数据的转置；数据的聚合；数据的拆分；数据的合并；数据的选择；数据的加权；数据的转换。

在本节中，将介绍几种常用的数据文件编辑方法。

1. 数据的排序

在进行数据的处理时，有时需要按照某个变量的取值重新排列各个个案在数据文件中的先后顺序。当读入原始数据文件后，可以按照下列步骤对数据文件中的个案进行排序操作。

（1）在主菜单中单击“Data”菜单项，展开下拉菜单。

（2）在下拉菜单中选择“Sort Cases”项，打开“Sort Cases”对话框，按照其取值大小进行排序的变量称为 Sort By 变量，在对话框中选择 Sort By 变量及排序方式（从大到小或从小到大），系统就可以根据 Sort By 变量数据的大小，按照要求对原文件中的数据进行排序。

排序时也可以选择多个 Sort By 变量，要求系统根据多个变量数据的大小对原文件中的个案进行排序。排序的结果与这些变量输入 Sort By 窗口的顺序有关，最先输入的称为第一 Sort By 变量，其后输入的分别为第二、第三……Sort By 变量。SPSS 将首先根据第一 Sort By 变量对数据进行排序，然后对这

一排序结果中大小相同的第一 Sort By 变量数据所对应的个案根据第二 Sort By 变量数据及排序方式进行排序，后面依此类推。

2．数据的转置

SPSS 可以将数据编辑器中的数据进行行、列互换，使原来的行变成列，原来的列变成行。这一功能可以通单击主菜单“Data”菜单项中的“Transpose”选项实现。

(1) 在主菜单中单击“Data”菜单项，展开下拉菜单；

(2) 在下拉菜单中选择“Transpose”项，打开“Transpose”对话框，将需要转置的变量名称输入“Transpose”对话框的左侧，执行即可。

转置后，原数据文件中的变量成为新数据文件中的个案，而原文件中的个案变成了新文件中的变量。要注意的是，在原数据文件中未被选定的变量在新文件中会丢失。另外，实施转置后，原数据文件中用户指定的缺失值将被转换为系统缺失值，若要重新获得这些值，需要重新进行定义。

3．数据的聚合

对数据进行聚合就是按照指定的分类变量取值进行分组，对每组观测量的各变量求描述统计量。

(1) 在主菜单中单击“Data”菜单项，展开下拉菜单；

(2) 在下拉菜单中选择“Aggregate”项，打开“Aggregate Data”对话框；

(3) 在左侧的源变量框中选择一个或多个变量作为分类变量（Break Variable）；

(4) 在左侧的源变量框中选择一个或多个变量作为聚合变量（Aggregate Variable），即要求 SPSS 对这些变量的值进行分类汇总；



(5) 单击“Aggregate Data”对话框中的“Name & Label”项，可以修改聚合变量的名称及变量标签。如果不进行修改，则系统对每个新变量赋予的变量名称均为对应的原始变量名称，后面加下划线和数字1；

(6) 单击“Aggregate Data”对话框中的“Function”项，可以打开“Aggregate Function”对话框，该对话框可以用来确定对聚合变量的描述内容。缺省的选项为Mean of Value，也就是求变量的平均值。

该对话框中的聚合函数有3类：第一类是对总体特征的描述，包括均值、总和等8项；第二类是对单侧区间的描述，有4种选择；第三类描述有上下界的区间特征，也有4个选项；

(7) 单击“Aggregate Data”对话框中的“Save number of cases in break group as variable”项，可以为新数据文件建立关于各分组个案数的新变量；

(8) 保存生成的新数据文件，可以建立一个新文件，也可以覆盖旧文件。

4．数据的选择

很多时候，对于某些特定的分析过程，只需要用到整个数据集的一部分，这就要从原始数据中选择出所需的数据。对此，SPSS提供了4种主要的方法。这4种方法可以从主菜单中“Data”菜单项中的“Select Cases”选项实现。打开“Select Cases”对话框，其中提供了5种数据的选择方式，缺省方式为选择全部个案。

(1) 根据逻辑关系表达式选择数据（If condition is satisfied）

单击“If condition is satisfied”按钮，打开If对话框，在列表框中输入逻辑关系表达式，其中数字、数学运算符、关系运算符可以通过键盘输入，也可以通过对话框中间的小键盘输

人，如果关系式中有函数，可以在“Function”列表框中进行选择。逻辑关系式输入完毕，回到主对话框运行，就可以实现数据选择。

例如，在数据集中，变量Q1表示居民对奥运会的关注程度，取值为1—5，取值越大，表示关注程度越高。而如果只想对关注程度高的居民作分析，在关系表达式窗口就可以输入“Q1 > 3”，利用这一逻辑表达式对数据进行选择。图10-17是选择后的运行结果，从中可以看到，Q1≤3的个案号都划上了斜线，表示这些个案已经被过滤掉了。

图10-17　用逻辑表达式选择数据的运行结果

（2）随机选取数据（Random Sample of cases）

单击“Random Sample of cases”按钮，打开“Random Sample”对话框，在此有两种方式来确定随机样本，第一种是输入选择数据个案数占总个案数的百分比，第二种是指定从A个个案中



随机抽取 B 个个案。两种选择方式的运行结果见图 10-18（a）、（b）。

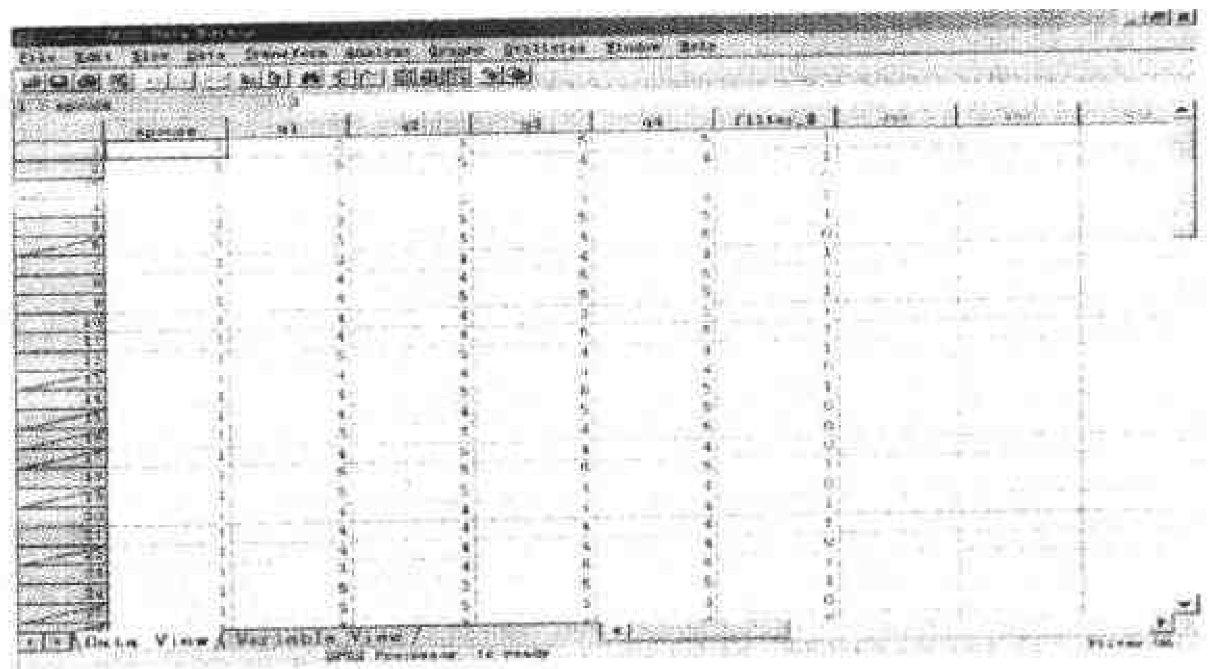


图 10-18（a） 随机选择 50% 个案的选择结果

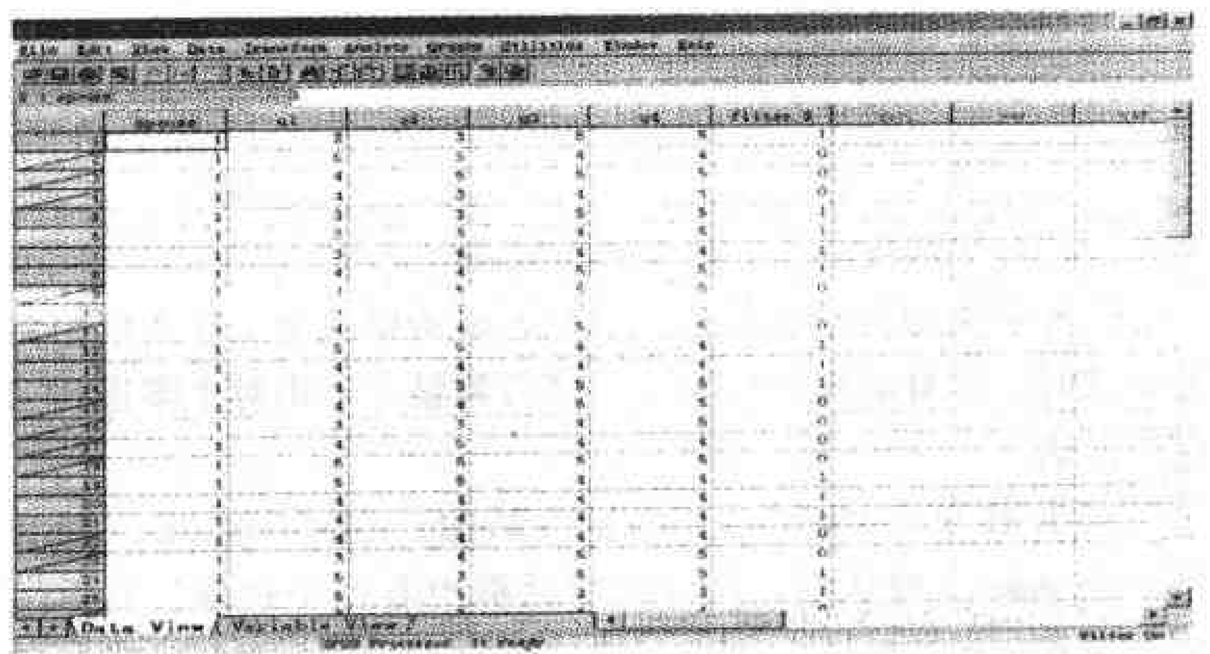


图 10-18（b） 60 个个案随机选择 30 个的结果

（3）在给定范围（日期、时间或个案号等）内选择数据（Based on time or case range）

单击“Based on time or case range”按钮，打开“Range”对话框，在First窗口输入起始个案号，在Last Case窗口输入终止个案号。例如，第一个输入10，第二个输入20，运行以后，原数据文件中只保留个案号为10~20的个案，其余个案均被滤掉了。

(4) 用过滤器变量选择数据（Use filter variable）

对应于在某变量上的取值为零或缺失的个案，如果你不打算对它们进行分析，可以用该变量作为过滤器变量，将在该变量上取值为零或缺失的个案剔除。

5. 数据的加权

加权是一种通过人为方法调节样本或数据大小的方法。所谓加权，就是给被加权对象乘上一个系数，当系数小于1时，即缩小被加权对象在整体中的影响；而系数大于1时，扩大被加权对象在整体中的影响。

加权通常可用于两种情况：

(1) 对变量取值进行加权

这可以通过“Compute”功能来实现，详见数据转换。

(2) 对个案进行加权（Weight Cases）

当所采集的样本不具有代表性，或需要改变子样本的相对重要性时，就有必要对样本大小进行调整，采用对个案进行加权的方法。

在主菜单“Data”中单击“Weight cases”选项，打开“Weight cases”对话框，在这个对话框中有两个选项，其中一个是缺省选项“Do not weight cases”，即不对个案进行加权；另一个选项为“Weight cases by”，选择该项，可以在对话框左边的变量列表框中选择一个变量，放入Frequency窗口中，则所选变量对应于每一个案的值便是对该个案的加权值。例如，表



10-5（a）是对样本性别的频数分析，可以看到其中男性的比例明显高于女性，为了改变这个比例关系，可以对个案加权，表 10-5（b）是对男性加权重为 1，女性权重为 2 之后的样本性别频数分析，加权后，女性的样本量扩大了一倍。

加权以后的数据文件可以进行保存，但是保存以后的数据文件重新打开以后，可以随时取消加权，使数据文件恢复到原数据文件的状况。

表 10-5　　加权前后对数据进行频数分析的结果

您的性别

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	男	125	67.9	67.9	67.9
	女	59	32.1	32.1	100.0
	Total	184	100.0	100.0	

(a)

您的性别

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	男	125	51.4	51.4	51.4
	女	118	48.6	48.6	100.0
	Total	243	100.0	100.0	

(b)

6. *数据的转换*

所谓数据转换，就是利用原有数据，通过某种转换关系生成新数据。下面介绍 SPSS 中比较常用的两种数据转换方法：

（1）利用“Compute”功能选项转换数据

利用“Compute”功能选项转换数据实际上是通过变量的计算实现数据转换。

单击主菜单“Transform”菜单项中的“Compute”选项，可以打开 Compute variables 对话框。在 Target 窗口中输入目标变量，可以是新变量，也可以替换原有变量。例如，输入新变量 Q1NEW，而对于计算目标变量的数学表达式则写在 Numeric Expression 窗口中。例如，输入的数学表达式为“Q1 + 20”，那么完整的计算关系为：Q1NEW = Q1 + 20。当然 B 应该是已经存在的变量。同时单击对话框中的 If…按钮，打开 If 对话框，可以通过逻辑表达式选择需要计算的变量值。例如，可以输入逻辑表达式“Q1 > 3”。图 10-19（a）和（b）显示了进行转换前后的数据情况。从中可以看到，对于满足 Q1 > 3 的个案，新变量 Q1NEW 比 Q1 变量的对应数据增加了 20。

图 10-19（a） 用 Compute 功能选项转换之前的数据



图 10-19（b）　用 Compute 功能选项转换之后的数据

（2）利用 Recode 功能选项转换数据

利用 Recode 功能选项可以直接更改变量的值。单击主菜单"Transform"中的 Recode 选项，即可打开次级菜单。Recode 功能根据更改变量值以后值的保存方式分为存入相同变量和存入不同变量两种。以一个例子来说明如何操作。

①将更改结果存入相同变量

图 10-20（a）是 SPSS 中的一些原始的数据，其中 Q1 取值有 2，3，4，5，尝试将取值为 4 的个案修改为 Q1 = 8。

在 Recode 次级菜单中单击 Into same variables 选项，打开 Recode into same variables 对话框。

在变量列表中选定并单击准备修改的变量，移到 variable 窗口中，本例为变量 Q1。

如果需要，可以利用 If…对话框加一些逻辑表达式，来控制只修改符合某些条件的变量。

单击"Old and New Values"按钮，打开 Old and New Values 对话框，输入变量旧的值和新的值，单击 Add，本例中旧值为 4，新值为 8。

图 10-20（b）为使用 Recode 功能项修改后的数据。对于变量 Q_1，实现了 4→8。

	q1	q2	q3	q4
1	3	3	5	5
2	5	5	4	4
3	4	5	5	5
4	4	3	4	4
5	3	3	5	5
6	3	3	4	5
7	3	4	4	4
8	4	4	5	5
9	4	5	5	5
10	4	4	3	5
11	4	4	5	5
12	5	5	4	4
13	4	4	4	4
14	4	3	5	5
15	4	4	5	5
16	3	3	4	5
17	4	5	4	5
18	5	5	5	5

图 10-20（a） SPSS 中的原始数据

	q1	q2	q3	q4
1	3	3	5	5
2	5	5	4	4
3	8	5	5	5
4	8	3	4	4
5	3	3	5	5
6	3	3	4	5
7	3	4	4	4
8	8	4	5	5
9	8	5	5	5
10	8	4	5	5
11	8	4	5	5
12	5	5	4	4
13	8	4	4	4
14	8	3	5	5
15	8	4	5	5
16	3	3	4	5
17	8	5	4	4
18	5	5	5	5

图 10-20（b） 修改后存入同一变量的数据结果

②将更改结果存入不同变量

在次级菜单中单击 Into different variables 选项，打开 Recode into different variables 对话框。与 Recode into same variables 对话框不同的是多了一个 Output Variable 窗口，用来输入输出变量的变量名称和标签。例如，在上例中，将 Q1 的值修改后，存入 Q1NEW 变量，其中 4 修改为 8，其余值不变。修改结果如图 10-20（c）所示。

图 10-20（c）　修改后存入不同变量的数据结果

在利用 Recode 功能选项修改变量取值时，尤其是将修改的结果存入相同变量，要特别小心，一旦操作失误，将可能无法恢复原始的数据。

第十一章　常用统计分析方法

第一节　统计概论

一、描述性统计学和推断性统计学

通过精确新闻调查收集的资料，需要进行统计分析，并对结果做出相应的解释。一般社会科学研究中所用的统计，可以分为描述性统计和推断性统计两种。统计研究过程的起点是数据，终点是探索到客观事物总体内在的数量规律性。如果我们收集到的是总体数据，使用描述统计方法，就可以达到探索数据内在规律性的目的；但是，经常的情况是，我们获得的数据只是研究总体的一部分数据，要探索到总体的规律性，必须应用概率论的理论并根据样本数据对总体做出科学的推断。

为了将调查样本中的信息向总体进行科学推断，必须首先对从样本中得到的大量数据资料进行整理、概括和计算，这就是描述统计的任务。描述统计是推断统计的基础。研究者一般使用描述统计描述或说明数据的意义。例如，北京市民的平均月收入是多少？农村地区的出生率和死亡率等，都是描述统



计。描述统计只就数据本身加以分析和解释，无法进行推论。

推断统计的目的是从样本信息来推断总体特征。一般来说，研究者进行抽样调查，他们真正关心的不是样本的信息，而是总体。研究者希望通过从样本得到的数据推断总体，必须借助推断性统计。例如，我们在北京市居民中抽取1000个样本，了解他们对北京市申办2008年奥运会的态度，而我们感兴趣的不是这1000人的看法，而是希望从这1000个样本的想法推断北京市全体市民对于这件事情的态度。这就是推断统计需要做的工作。

二、单变量分析、两变量分析和多变量分析

在对收集到的数据进行统计分析时，无论采用描述统计还是推断统计，都可以按照所解释的变量数目分为单变量分析、两变量分析与多变量分析。

如果在某次数据分析时，只分析或解释一个变量，则称为**单变量分析**。例如，盖洛普公司所做的关于北京市民对于北京申办2008年奥运会态度的民意调查，共访问了1626位北京市民。从分析结果来看，94.9%的北京市民表示支持，4.4%表示无所谓，0.3%表示反对。在这个分析中，由于只涉及到是否支持这样一个变量，所以是单变量分析。

如果在统计分析中一次涉及两个变量，即称为两变量分析。比如，在上面关于北京市民对于北京申办2008年奥运会态度的民意调查中，如果我们进一步分析不同性别的市民对申办奥运会的支持程度，发现样本中男性对于北京申办奥运会表示支持的为94%，而女性为96%。这个分析涉及“性别”和“支持与否”两个变量，因此是两变量分析。

由于社会现象是错综复杂的，两个变量之间的关系，很多

时候还会受到其他变量的影响。因此，在对社会现象进行分析时，除了进行单变量和两个变量的统计分析以外，可能还需要进行多变量分析。所谓多变量分析是指每次分析可能同时涉及3个或3个以上的变量。如果我们不仅考虑性别的影响，同时希望了解学历对于北京市民对申办奥运会态度的影响，那么就需要同时分析“性别”、“学历”和“支持与否”3个变量之间的关系了，教育程度高的男性如何，教育程度高的女性如何，教育程度低的男性如何，教育程度低的女性如何。这就是一个多变量分析。

随着电子计算机技术的发展，过去被认为是具有数字高难度的多元统计分析方法，已经被越来越广泛地应用于实际。但是，由于多变量分析涉及比较复杂的统计方法，已经超出了本书的范围，这里将不作介绍，而只介绍单变量分析和两变量分析。

第二节　描述性统计

一、统计量

对于抽取的样本，常常需要用一些概括性的数量对它进行描述，这种概括性的量就是所谓的“统计量”。

例如，在前面提到的北京市民对北京申办2008年奥运会的民意调查，1626个样本中对是否支持北京申办奥运会的回答分别是：“支持”、“无所谓”、“不支持”等。对这个样本的一个适当的概括性数量便是统计量P：“支持”北京申办奥运会的人数占样本的比例。这个统计量P可以这样计算，用样本中表示“支持”的人数除以样本量，就可以得到P=0.949。

当然，如果我们在北京市再次独立地抽取一个n=1626的



样本，调查同样的问题，由于抽样的不确定性，回答“支持”的人数比例可能就不再是0.949。也就是说，统计量的值是随样本而变的，是样本的函数，而且其中不包含任何未知的参数。所以对于统计量的描述性定义为：统计量是一个不包含任何未知参数的样本函数。

二、频数（Frequency）分布

通常一个精确新闻的研究包含很多变量和许多研究对象，因此在进行数据分析时就要面对大量的原始数据。如抽取600个样本，问卷中涉及30个变量，就会有1.8万个原始数据。这样大量的数据，表面上往往是杂乱而没有规律的，所以首先要对数据进行组织整理，以便能对数据的全貌有个初步的了解。计算频数分布是最常用的定量而直观地描述数据的手段之一，是最基本的统计分析方法。

频数是指变量每一个取值出现的次数。

1. 离散变量的频数

例如，在一项关于人们对于新世纪看法的研究中，有一道题目是了解被调查者认为21世纪人们信息的最主要来源是什么（X）？这个变量X的取值是离散的：1＝“电视”，2＝“网络”，3＝“报纸”，4＝“广播”，5＝“报纸”，6＝“其他”，研究的样本量n＝600。如果研究者想知道在600个样本中，变量X各个取值出现的次数是多少，就可以采用频数分析。

2. 用SPSS进行频数分析

SPSS的Frequencies过程可以实现计算变量的频数分布，产生详细的频数表：

• 打开Analyze菜单，单击Descriptive Statistics子菜单下的Frequencies菜单项，打开Frequencies对话框；

• 该对话框左侧为变量名列表框，其中列出了当前数据文件中所有变量的变量名称；

• 对话框的右侧为 Variable（s）列表框，用于选定需要进行频数分析的变量，只需在变量名列表框中选定要分析的变量名称，将其移入 Variable（s）列表框，SPSS 将对选定变量进行频数分析；

• 选择 Display frequency table 核选框，将在查看器中显示频数分析表，否则只显示直方图；

• 通过该对话框中的 Statistics 可以定义需要计算的其他描述统计量；Charts 按钮可以设定所做的统计图；Format 按钮可以定义输出频数表的格式。

在 SPSS 中按照以上步骤计算上例中变量 X 的频数分布的结果如表 11-1 所示：

表 11-1　SPSS 输出的频数表

Statistics

21 世纪人们信息的主要来源是什么？

N	Valid	600
	Missing	0

(a)

21 世纪人们信息的主要来源是什么？

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	电视	127	21.2	21.2	21.2
	网络	380	63.3	63.3	84.5
	报纸	50	8.3	8.3	92.8
	广播	6	1.0	1.0	93.8
	杂志	6	1.0	1.0	94.8
	其他	31	5.2	5.2	100.0
	Total	600	100.0	100.0	

(b)



在表 11-1（a）中，输出的是数据的基本描述统计量，其中可以看出参与本次计算的有效样本数 N 为 600 个，缺失值为 0 个。如果需要计算其他的统计量，可以在 Frequencies 对话框中利用 Statistics 按钮来设定。

在表 11-1（b）中 Frequency 即是频数分析，Total 代表样本总数为 600，其中认为“电视”是 21 世纪人们信息主要来源的有 127 人，认为“网络”是信息主要来源的有 380 人，认为“报纸”是信息主要来源的有 50 人，另外分别有 6 人认为“广播”或“杂志”是信息主要来源，31 人认为信息主要来源于除以上媒介的其他渠道。

在表 11-1（b）中，还包含有百分比（Percent）、有效百分比（Valid Percent）和累积百分比（Cumulative Percent）。我们在后面会对这几个量进行说明。

3．连续变量的频数

如果在研究中测量的是连续型变量，如，北京市民的家庭收入，那么得到的变量 Y 就是一个连续型变量，它的取值可能是 1500 元，可能是 3000 元，也可能是这两者之间的任意一个数值。此时计算变量 Y 的某一特定取值所对应的频数是多少就不再有意义了，这种情况下，可以采用分组的形式对连续型变量进行整理。

例如，我们在北京市抽取了 n = 200 个样本，询问这些样本家庭平均月收入是多少元（变量 Y)？数据收集回来后，可以先使用前面介绍的 SPSS 中 Recode 命令对数据进行重新编码，将家庭月收入在 1000 元以下的编码为 1，收入在 1001 ~ 2000 元之间的编码为 2……然后计算分组后的家庭月收入频数分布，见表 11-2：

表 11-2　居民家庭月收入的频数分布

Statistics

家庭平均月收入

N	Valid	200
	Missing	0

家庭平均月收入

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1000 元以下	27	13.5	13.5	13.5
	1001 ~ 2000 元	60	30.0	30.0	43.5
	2001 ~ 3000 元	74	37.0	37.0	80.5
	3001 ~ 4000 元	14	7.0	7.0	87.5
	4000 元以上	25	12.5	12.5	100.0
	Total	200	100.0	100.0	

这样从表中就可以看出，多数家庭的月收入在 1001 ~ 3000 元之间。

如果需要对连续型的变量作频数分布分析，最好先找出变量的取值范围，比如，上例中，家庭月收入的取值范围是 500 元到 2.5 万元。之后决定分组数和组距，注意组数一般不要太多或太少，组距一般取整数为好。按照决定的组数和组距分组后，统计观测数据落入各组的频数。

三、频率和百分数

1．频率

所谓频率是将频数除以样本总数，即 $P = f/n$，其中 f 为变量某个取值出现的频数，n 为所研究的样本数。例如，在表 11-1 中，认为“网络”将是 21 世纪信息主要来源的样本频率



为：$P = f/n = 380/600 = 0.633$。

有些时候，频数分析不能说明问题，需要计算频率，特别是在对不同的样本数进行研究时，下面是一个简单的例子：

表 11-3　男性样本的学历

学历	f	P
初中及以下	110	0.200
高中、中专	288	0.524
大专及以上	152	0.276
合计	550	1.000

(a)

表 11-3　女性样本的学历

学历	f	P
初中及以下	50	0.104
高中、中专	295	0.615
大专及以上	135	0.281
合计	480	1.000

(b)

由表 11-3（a）和表 11-3（b）可见，大专及以上样本占男女样本总数的比例分别是 0.276 和 0.281。如果只是看频数，男性中大专及以上人数比女性多，但这种比较是存在问题的，因为两个样本的总数不同。这时频数就不能说明问题，而需要使用频率来比较。因为频率计算使得两个样本的总数变成同一个基数 1，这样就可以进行比较了。从上面计算的频率值，可见男性和女性中大专及以上人所占的比例相差甚少。

2．百分数

在对数据进行统计分析时，经常使用百分数代替频率。百分数也是精确新闻报道最为常用的统计量。计算百分数的公式是：$P = f/n \times 100$，即将频率放大 100 倍。

3. 用SPSS计算百分数

在利用SPSS对变量进行频数分析的同时，SPSS也会给出频数对应的百分数，见表11-1（b）。在该表中，Percent列即为相应频数占样本总数的百分数，例如，21.2是指认为“电视”将是21世纪信息主要来源的人（127）占样本总数（600）的21.2%，63.3是指认为“网络”将是21世纪信息主要来源的人（380）占样本总数（600）的63.3%……

表11-1中的Valid Percent是有效百分数，有效百分数是指SPSS真正计算的百分数。如果样本在该变量上的数据有缺失个案，有效百分数在计算时，会从样本总数中将缺失个案剔除，再计算相应频数占实际该问题样本数的百分数。在表11-1（b）中，由于没有缺失个案，因此百分数与有效百分数列的数值完全相同。假如这个变量有5个人没有回答，即有5个缺失个案，此时SPSS在计算有效百分数时，将会自动去掉5个缺失个案，回答样本总数变为595人，而认为“网络”将是21世纪信息主要来源的人(380)的有效百分数为380/595≈0.639，有效百分数与百分数就不再相同了。

Cumulative Percent是累积百分数，即各个频数所占百分数的累积总和，仍然以表11-1（b）为例，Cumulative Percent列中84.5就代表认为“电视”将是21世纪信息主要来源的人(21.2%)，再加上认为“网络”将是21世纪信息主要来源的人（63.3%），共占样本总数的84.5%。如果再加上“报纸”、“广播”、“杂志”和“其他”项目的百分数总和将达到100%。

在推断性统计中，百分数也常用于估计相同的项目在调查总体中所占的比例。百分数虽然简单，但在使用中极易出现各种问题，所以要特别注意。

首先，计算百分数时使用的基数不能太小。例如，在一项居



民使用互联网情况的研究中，在分析不同年龄的人使用互联网的情况时，发现65岁以上老年人中使用互联网的比例是最高的，达到50%。事实上，样本中只有4个65岁以上的老人，其中恰巧有两位使用过互联网。可见用这么小的基数计算的百分数是毫无意义的，如果回避了对基数的说明，更会产生极大的误导。

其次，在用百分数的大小进行比较时，还应当适当考虑绝对数的大小。百分数的大小提供了一个相对的概念，但是当计算基数不同时，有时绝对数的大小也应该加以说明。

四、变量取值集中趋势的描述

对变量取值集中趋势的描述，主要就是寻找一个数值来表示变量的数据集中情况。必要的时候，可以根据这个代表值（典型值）来估计或预测每个个案的数值。

最常用于变量取值集中趋势描述的3个统计量为：众数、中位数和平均数。

1. 众数（Mode）

一个变量的众数定义为频数最多的变量值。例如，在表11-1中，变量X的众数是2，也就是21世纪信息主要来源的众数是“网络”（f=380）。

众数很容易求，一般都可以直接看出。众数适用于描述定类变量的数据，当然也可以用来分析定序或定距变量的数据。例如，表11-2显示，北京市居民家庭平均月收入的众数是2001～3000元（f=74）。对于定距变量，也可以用数据分组后的组中值来表示众数的大小。

变量的众数很不稳定，并不是一个描述变量数据集中趋势的很好的统计量。它常常依赖于数据的分组情况，分组改变的话，众数可能就会有较大的变化。而且众数有可能不是惟一

的，例如，变量的频数分布可能出现两个峰值。

2．中位数（Median）

中位数就是在一个序列的中央位置的值，是第50百分位数点上的值，一般用Md表示，即取值高于此值的有50%的样本，低于此值的样本也有50%。计算中位数分两种情况，一种是使用原始数据，另一种是用分组数据。

（1）根据原始数据求中位数

计算方法是将各个个案由低至高排列，居序列中央位置的样本的值就是中位数。例如，甲村5户人家的人口数由少至多排列如下：2、3、4、6、7。这个序列的中位数就是4，其中有两户多于4人，两户少于4人。根据原始数据计算中位数的公式是：

Md位置＝（n＋1）/2

其中n为样本数。上例中，n＝5，Md位置＝（5＋1）/2＝3，而序列中第三个位置的值就是4，因此Md＝4。

上例中n为奇数，如果n为偶数，通常的做法，是将位于最中央的两个数值的平均值作为中位数。例如，乙村8户居民的人口数由少至多分别是：2、4、4、5、7、8、9、9。则按照以上公式，Md位置＝（8＋1）/2＝4.5，这表示中位数的位置落在序列上的第四户与第五户之间，即Md＝（5＋7）/2＝6。

（2）用分组数据求中位数

表11-4　北京市居民家庭月收入的分组情况

家庭平均月收入	频数 f	累积频数 cf
1000元以下	27	27
1000～2000元	60	87
2000～3000元	74	161
3000～4000元	14	175
4000元以上	25	200
合计	200	



前面我们曾将北京市居民的家庭平均月收入分为5组，如表11-4所示。计算这样分组数据的中位数，首先需要计算各组的累积频数，继而求出中位数的位置：Md位置＝（200＋1）/2＝100.5。根据累积频数的分布，第100.5的位置在2001～3000元组内，计算公式如下：

$$Md = L + (n/2 - cf)/f \times w$$

其中，L＝中位数所在组的真实下限；

f＝中位数所在组的频数；

w＝中位数所在组的组距；

cf＝低于中位数所在组真实下限的累积频数

n＝全部样本数

根据这个公式来计算表11-4分组数据的中位数：

$$Md = 2000.5 + (200/2 - 87)/74 \times 1000 = 2176$$

这个中位数告诉我们有一半的样本家庭平均月收入在2176元之下，另一半在2176元之上。

由于中位数的定义依赖于数据的大小顺序，因此它不适合于定类变量。对于定序和定距变量，特别是对于数据分布不很规则的情况，中位数是比较理想的统计量之一，同时中位数对极端值不敏感。

3．平均数（Mean）

平均数也是精确新闻报道常用的统计分析方法。平均数适用范围极广。例如，研究居民收看电视的时间、男性和女性的工资差异等等问题，都需要采用平均数分析。平均值的计算方法，也分为两种情况，即使用原始数据和使用分组数据。

（1）使用原始数据计算平均数

计算公式是：平均数 $\overline{X} = \sum x/n$

其中 $\sum x$ 表示各个个案在该变量上的取值之和，n 表示全

部个案数目。比如，调查5个小学生的语文课成绩分别是75、87、78、91、83，根据公式计算5个学生的平均成绩为：

$\bar{X}=(75+87+78+91+83)=82.8$

(2) 根据分组数据计算平均数

计算公式是：平均数 $\bar{X}=(\sum f x_m)/n$

其中，f 表示每组的频数，x_m 表示组中值，n 表示全部个案数目。例如，调查170位青年人一年来阅读小说的数目，表11-5是得到的数据。

表11-5　青年人阅读小说的数目

小说数目	频数 f	组中值 X_m	fx_m
2～4	20	3	60
5～7	40	6	240
8～10	50	9	450
11～13	30	12	360
14～16	20	15	300
17～19	10	18	180
合计	170		1590

按照公式计算青年人阅读小说的平均数为：

$$\bar{X}=1590/170=9.4$$

即平均来说，每名青年人一年阅读过9.4本小说。

平均数是最典型的也是最常用的统计量，适用于定距和定比率变量。平均数可以看成是数据的“平衡点”或“中心”位置所在。由于平均数的计算需要用到所有的样本数据，因此与中位数和众数相比，它所含的信息量最大。但是平均数对极端值比较敏感。

4．用SPSS计算众数、中位数和平均数

(1) 使用SPSS中Frequencies过程计算



在计算变量的频数分布的同时，可以利用 Frequencies 对话框中的 Statistics 按钮，定义输出众数、中位数和平均数。使用这种方法计算的表 11-2 所示北京市民家庭月收入的众数、中位数和平均数如表 11-6 所示。

表 11-6　北京市民家庭月收入的众数、中位数和平均数

Statistics

家庭月收入

N　Valid	200
Missing	0
Mean	3644.85
Median	2600.00
Mode	2000

由此可见，北京市居民家庭月收入的众数是 2000 元，中位数是 2600 元，平均数是 3644.85 元。

(2) 使用 SPSS 中 Means 过程计算分组平均数和中位数

在进行数据分析时，有的时候会需要比较不同子样本的平均数等统计指标，这可以通过 SPSS 中的 Means 过程实现。Means 过程对准备比较的各组计算描述统计量。

• 打开 Analyze 菜单，单击 Compare Means 子菜单下的 Means 菜单项，打开 Means 对话框；

• 对话框左侧为变量名列表框，其中列出了当前数据文件中所有变量的变量名称；

• 对话框的右侧 Dependent list 框，用于选定需要进行分析的变量。例如，将代表北京市居民家庭月收入的变量放入其中。

• 对话框的右侧 Independent list 框，用于选定分组变量。例如，将代表被调查者的学历的变量放入其中。

• 单击 Options 按钮，可选择需要的描述统计量。例如，选择输出变量的分组平均数和中位数。

在 SPSS 中按照以上步骤计算的结果如表 11-7 所示：

表 11-7 不同学历的北京市居民家庭月收入平均数和中位数

Report

家庭月收入

学历	Mean	N	Median
初中及以下	2197.83	46	2000.00
高中、中专	2669.23	65	2600.00
大专及以上	5105.28	89	3000.00
Total	3644.85	200	2600.00

从这个可以得到按照不同学历分组的北京市民家庭月收入的平均数和中位数。其中初中及以下学历组的家庭月收入平均为 2197.83 元，中位数为 2000 元；高中、中专组的平均数为 2669.23 元，中位数为 2600 元；大专及以上组的平均数为 5105.28 元，中位数为 3000 元。

在描述变量取值集中趋势的统计量众数、中位数和平均数中，平均数最为稳定，众数稳定性最差。但是平均数受极端值影响较大，而中位数对极端值不那么敏感，因此，当数据分布比较有规则，不存在极端值时，用平均数描述变量的集中趋势比较好，而在有极端值时则用中位数更为合适。例如，有些地区的家庭贫富极为悬殊，计算平均收入就没有什么意义，可以改用中位数表示地区的经济水平。

五、变量取值离散趋势的描述

如果变量取值的差异很大，众数、中位数和平均数的代表



性就会降低。因此，我们不仅要描述变量取值的集中趋势，而且要测量变量取值的离散趋势，也就是说，还要用一些统计量描述变量取值围绕中心向两个方向伸展的程度。

1．极差（Range）

极差就是变量取值的最大值和最小值之间的距离。

极差 R = 变量最大观测值 - 变量最小观测值

它指示了变量取值的伸展范围。例如，北京市居民家庭平均月收入的调查中，收入最大值为 50000 元，最小值为 500 元，极差 R = 50000 - 500 = 49500 元。

极差只告诉我们变量的取值范围，至于中间的数据是如何变化的则不得而知，而且极差受极端值的影响非常大，有时是不可靠的，所以并不常用。

2．四分位数间距（Interquartile Range）

(1) 定义

四分位数间距也叫四分位差。通常将数据按从小到大的顺序排列，然后用 3 个四分位数点 Q_1，Q_2，Q_3 将其分为四等份，每等份包含 25%的个案（见图 11-1）。其中 Q_1 是第 25 百分位数点或称低四分位数点；Q_2 是第 50 百分位数点，即中位数；Q_3 是第 75 百分位数点或称高四分位数点。四分位数间距定义为：

$$IQR = 高四分位数点 - 低四分位数点 = Q_3 - Q_1$$

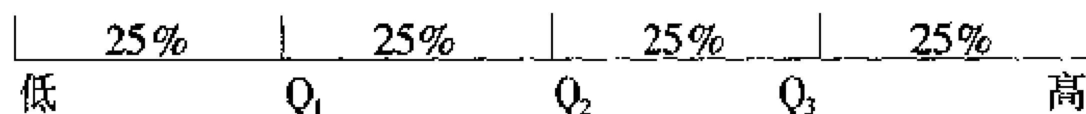


图 11-1　四分位数

很显然，计算四分位数间距时，要先求出 Q1 和 Q3 的位置，然后计算在这两个位置上的值的差异。计算方法与中位数的计算类似，分为对原始数据计算和对分组数据计算两种，这里只介绍对原始数据计算四分位数间距的做法。对于分组数据

的计算，有兴趣可以参看有关书籍。

(2) 根据原始数据计算四分位数间距

根据原始数据计算四分位数位置的公式是：

Q_1 位置 = (n+1) /4

Q_3 位置 = 3× (n+1) /4

其中 n 为样本数。例如，我们在某个农村抽了 8 户人家，每户人家的人口数依次为：

2，3，4，7，9，10，12，12

因为 Md 位置 = (8+1) /2 = 4.5　Md = (7+9) /2 = 8

Q_1 位置 = (8+1) /4 = 2.25　Q_1 = 3 + 0.25 × (4 − 3) = 3.25

Q_3 位置 = 3 × (8+1) /4 = 6.75　Q_3 = 10 + 0.75 × (12 − 10) = 11.5

所以　$Q = Q_3 - Q_1 = 11.5 - 3.25 = 8.25$

四分位数间距虽然也是两个值之差，但是这两个值之差代表了中间 50%数据部分的长度，所以比极差更较好地描述数据的分布。如果 IQR 比较小，则说明数据比较集中在中位数附近；反之则比较分散。IQR 常常和中位数一起用来描述定序变量的分布，也可以描述定距变量的分布。

3．*方差和标准差*

分析定距变量取值最常用的方法是计算其方差或标准差。*方差* S^2 或*标准差* S 是衡量变量所有取值偏离其平均数 $\overline{X}$ 的程度的统计量。对原始数据计算方差的公式为：

$$方差\ S^2 = \frac{1}{(n-1)}\sum (X - \overline{X})^2$$

其中 n 为样本数，X 为变量观测值，$\overline{X}$ 为变量取值的平均数。对分组数据，方差的计算公式为：



方差 $S^2=\frac{1}{(n-1)}\sum(X-\bar{X})^2f$

其中 n 为样本数，X 为变量每组观测值的组中值，$\bar{X}$ 为变量取值的平均数，f 为每组的频数。

定义方差的正的平方根为标准差：

标准差 $S=\sqrt{S^2}$

方差和标准差反映的是变量取值对其平均数的某种离散程度，方差或标准差较小时，数据一定是比较集中在平均数附近；反之则是比较分散的。

4. 使用SPSS计算极差，四分位数和方差（标准差）

（1）使用SPSS中Frequencies过程计算

在计算变量的频数分布的同时，可以利用Frequencies对话框中的Statistics按钮，定义输出四分位数（Quartiles）、极差（Range）、方差（Variance）和标准差（Std.Deviation）。使用这种方法计算的表11-2所示北京市民家庭月收入的四分位数、极差、方差和标准差如表11-8所示。

表11-8　北京市民家庭月收入的四分位数、极差、方差和标准差

Statistics

家庭月收入

N	Valid	200
	Missing	0
Mean		3644.85
Std.Deviation		6804.05
Variance		46295063
Range		69500
Percentiles	25	2000.000
	50	2600.00
	75	3000.00

由上表可见，北京市居民家庭月收入平均数为3644.85元；分布极差为69500，数值很大，说明存在极端值；因为极

端值的存在，所以标准差（方差）的数值也很大，说明数据偏离平均数的程度较大；而由四分位点可以计算出四分位数间距是 1000 元，说明有 50% 的数据比较集中在中位数附近。

上例也说明了标准差（方差）虽然稳定，但受极端值影响比较大，极差对极端值也比较敏感，而四分位数间距虽然稳定性不如标准差，但对极端值并不敏感。

(2)使用 SPSS 的 Descriptives 过程计算极差、方差和标准差

• 打开 Analyze 菜单，单击 Descriptive Statistics 子菜单下的 Descriptives 菜单项，打开 Descriptives 对话框；

• 对话框左侧为变量名列表框，其中列出了当前数据文件中所有变量的变量名称；

• 对话框的右侧为 Variable（s）列表框，用于选定需要进行描述的变量，例如，将代表北京市居民家庭月收入的变量放入其中；

• 单击 Options 按钮，可选择需要的描述统计量。例如，选择输出变量的平均数、极差、方差和标准差。

在 SPSS 中按照以上步骤计算的结果如表 11-9 所示：

表 11-9　北京市居民家庭月收入平均数、极差、方差和标准差

Descriptive Statistics

	N	Range	Mean	Std. Deviation	Variance
家庭月收入	200	69500	3644.85	6804.05	46295063
Valid N (listwise)	200				

比较表 11-9 和表 11-8，可以看到利用 SPSS 两个不同的过程计算，得到了相同的结果。

六、正态分布与标准值

对于连续型变量，其分布形状有很多种，其中最重要的就

是正态分布曲线（Normal Curve）。此曲线呈钟型，其数学表达式为：

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-(X-\bar{X})^2/2S^2}$$

其中 X：变量的取值；f（x）：该变量值出现的概率；S：标准差；$\bar{X}$：平均数。

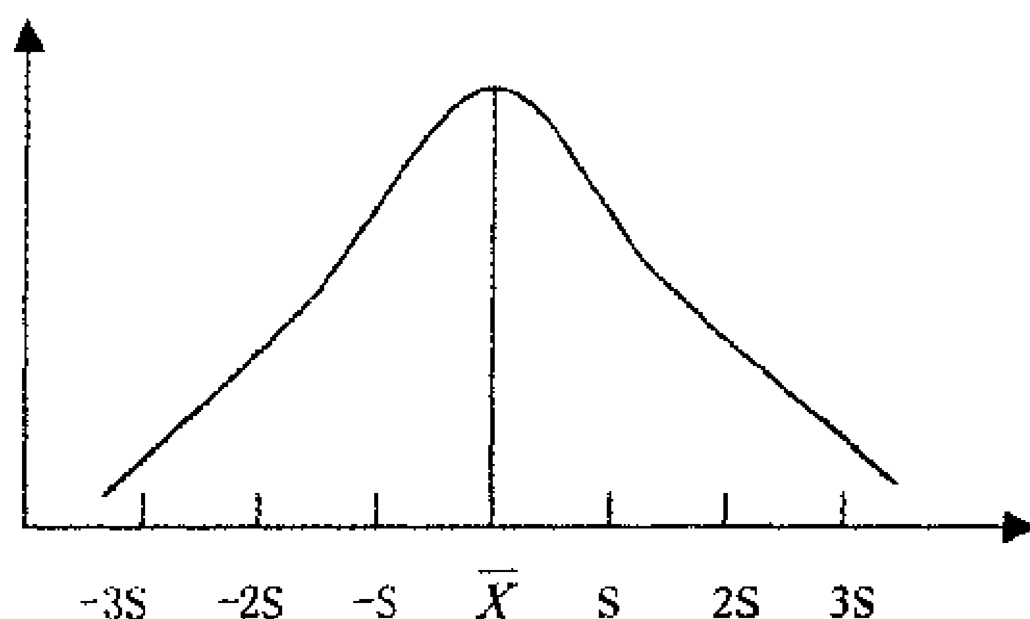


图 11-2 正态分布曲线

正态分布具有单峰和对称的特质，因此众数、中位数和平均数是相同的。变量取值 X 在某一范围内的概率，就等于相应范围内正态曲线下的面积。根据运算，在平均数两旁正负一个标准差范围内所包括的面积，约占全部面积的 68.26%，也就是说，约有 68.26% 的个案取值是在 $\bar{X}+S$ 和 $\bar{X}-S$ 之间；而有 95.46% 的个案是在 $\bar{X}+2S$ 和 $\bar{X}-2S$ 之间；99.37% 的个案在 $\bar{X}+3S$ 和 $\bar{X}-3S$ 之间。

例如，我们研究中学生的身高，身高一般是服从正态分布的，如果研究中发现中学生的身高平均为 160 厘米，标准差是

10厘米，那么我们就可以推断，中学生中，有68%的身高在150厘米到170厘米（160±10）之间；有95%的身高在140厘米和180厘米之间。

变量平均数等于0，标准差等于1的正态分布称为标准正态分布（Standard normal distribution）。所有正态分布都可以转化为标准正态分布，方法是以标准差作为计算单位，而放弃原来变量的度量单位，这样做的好处是可以使正态分布标准化，不受变量的度量单位所影响。

如果正态分布是以标准差（S）为单位，则每个变量值就变为：$Z=(X-\overline{X})/S$，其中Z称为标准值，代表每个变量值在标准正态分布上的数值。这个过程称为标准化过程。

正态分布是非常重要的统计分布，例如，人的智商、身高、体重，还有人对某种态度的得分，都近似服从正态分布。另外，许多统计方法都是建立在假设所抽取的样本变量取值呈现正态分布的基础之上的。

第三节　参数估计和假设检验

一、常用的抽样分布

精确新闻的研究主要目的是通过对样本数据的分析，而实现对研究总体的推断。例如，我们在北京市抽取200个样本，研究他们的家庭月收入，其目的主要是了解北京市居民总体的家庭月收入水平。

为了通过调查的样本对目标总体进行统计推断，必须知道一些常用统计量的分布，称之为抽样分布。

举例来说，要研究某市所有青年（总体）的每月用于休闲



活动费用情况，假设其平均数是 $\mu=70$ 元。现在从这个总体中不断地抽取随机样本，而每次抽取的样本大小（n）都是相同的，如果在每次抽取样本中都计算其平均数（$\bar{X}i$），就会得到许多不同的平均数，如 65.2 元，70.0 元，72.4 元，71.8 元，70.0 元，68.2 元，49.0 元，58.6 元等等，抽取的次数多了，其中当然有些平均数是重复出现的。如果把每个平均数的出现次数加起来，就会得到样本平均数的频数分布情况，有些值出现很多次，有些则比较少出现。这个频数分布情况就称为抽样分布。

1．样本平均数的分布

假定变量 X 的平均数为 μ、标准差为 σ，这里所说的平均数和标准差是指研究总体中变量 X 的平均数和标准差，$\bar{X}$ 为从这个研究总体中抽取的一个非常简单随机样本的样本平均数。通过不断的重复抽样实验，可以验证以下结论：

- $\bar{X}$ 的期望 $=E(\bar{X})=\mu$
- $\bar{X}$ 的标准误差 $=SE(\bar{X})=\sigma/\sqrt{n}$
- 如果变量 X 服从正态分布，那么样本平均数 $\bar{X}$ 也服从正态分布
- 如果 X 不服从正态分布，那么当样本量较大时，$\bar{X}$ 也近似地服从正态分布。

以上几点可以概括为正态近似定理如下：

在样本量为 n 的非常简单随机样本中，样本均值 $\bar{X}$ 以 $\sigma/\sqrt{n}$ 的标准误差围绕着总体均值 μ 波动。随着 n 的增大，$\bar{X}$ 的分布围绕着 μ 的波动也就越来越小，它也越来越接近正态。

上述定理告诉我们，如果样本是从总体中随机抽取的，那么平均来说，用样本均值可以正确地估计总体均值，估计的精度取决于所抽取的样本量。样本量越大，对总体平均数的估计

也就越准确。根据正态分布的理论可知，如果从总体中重复抽取许多容量为 n 的非常简单随机样本，那么：

有 90%左右的样本平均数会落入 $\mu \pm 1.64SE$ 的范围之内；

有 95%左右的样本平均数会落入 $\mu \pm 1.96SE$ 的范围之内；

有 99%左右的样本平均数会落入 $\mu \pm 2.58SE$ 的范围之内。

2. 样本比例的分布

假设变量 X 在研究总体中取某个值的比例为 π，可以计算总体标准差为：

$\sigma = \sqrt{\pi(1-\pi)}$，若 P 为来自该总体的一个非常简单随机样本的样本比例，与样本平均数的抽样分布研究方法类似，通过不断的重复抽样，可以得到以下结论：

- P 的期望 = $E(P) = \pi$
- P 的标准误差 = $SE(P) = \sigma/\sqrt{n} = \sqrt{\pi(1-\pi)/n}$

由此样本比例的正态近似定理为：

在容量为 n 的非常简单随机样本中，样本比例 P 以 $\sqrt{\pi(1-\pi)/n}$ 的标准误差围绕着总体比例 π 波动。随着 n 的增大，P 的分布也就围绕着 π 波动得越来越小，它也就越来越接近于正态。

由上述定理可知，如果样本是从总体中随机地抽取的，那么平均来说，用样本的比例可以正确地估计总体比例，估计的精度取决于所抽取的样本量。样本量越大，对总体比例的估计也就越准确。如果样本量足够大，那么样本比例 P 的抽样分布会很接近正态分布。根据正态分布的理论可知，如果从总体中重复抽取许多容量为 n 的非常简单随机样本，那么：

有 90%左右的样本比例值会落入 $\pi \pm 1.64SE$ 的范围之内；

有 95%左右的样本比例值会落入 $\pi \pm 1.96SE$ 的范围之内；



有 99%左右的样本比例值平均数会落人 $\pi \pm 2.58SE$ 的范围之内。

前面我们提到的推断性统计，就是根据抽样分布的原理进行的。只要在研究中采用随机抽样的方法，就可以根据抽样分布的理论，以样本的统计量来推断总体的情况。

推断统计一般分为两大类，参数估计和假设检验。所谓参数估计，就是根据一个随机样本的统计量来估计总体参数是多少。例如，由样本数据计算的青年人每个月用于休闲活动的花费是 72.5 元，然后根据这个数值，对所有该市青年（总体）中的花费平均数做估计。所以说参数估计的方法是先看样本情况，然后推断总体情况。而假设检验，首先是假设总体的情况是怎样的，然后以一个随机样本的统计值来检验这个假设是否正确。例如，对于刚才的问题，我们在研究之处，先假设该市青年在休闲活动中的平均花费是 70 元，然后抽样，通过计算样本的平均花费来检验原先的假设是否正确。所以假设检验是先假设总体的情况，然后进行抽样，通过样本来检验原来对于总体的假设是否正确。

二、参数估计（Parameter estimation）

用样本的统计量估计总体的参数值，有两大类做法，一种是点估计，另一种是区间估计。

1. 点估计

所谓点估计，就是用一个最适当的样本统计量来代表总体的参数。例如，要研究北京市民对于北京市申办 2008 年奥运会的态度，从北京市抽取了一个随机样本，而计算出样本中表示支持的比例为 94.8%，那么我们就可以说，北京市民中有 94.8%的人支持北京申办奥运会。如果我们研究的是北京市民

家庭月收入问题，就可以把样本的家庭月收入平均数作为对北京市民总体的估计值。

一般来说，如果样本越大，且抽样方法越严谨，这种估计就越可信。但是无论如何，抽样误差总是难免的，而且由点估计无法得到估计值的可信程度。因此，更多的时候是采用区间估计。

2．区间估计

所谓区间估计，就是用两个数组成的区间估计参数值。至于取多大的区间，取决于在估计时所要求的置信程度（Level of confidence）是多少。在样本大小相同的情况下，要求的置信度越大，则区间就越宽。这个区间，就称为置信区间（Confidence Interval）。例如，从样本中计算出青年人中赞成某项新政策的比例是 60%，在由此估计总体的赞成比例时，如果要求置信度为 95%，则通过计算，置信区间应该是［55%，65%］。假如要求置信度为 99%，则置信区间扩大为［53%，67%］。

在采用区间估计时，可以自己决定置信度，但用的比较多的是 95%或 99%。下面介绍几种统计量的置信区间计算方法。

（1）总体平均数的区间估计

设置信度为 95%，总体平均数的置信区间为：

$$\mu=\overline{X}\pm 1.96SE=\overline{X}\pm 1.96\ (\frac{S}{\sqrt{n}})$$

置信度为 99%，总体平均数的置信区间为：

$$\mu=\overline{X}\pm 2.58SE=\overline{X}\pm 2.58\ (\frac{S}{\sqrt{n}})$$

其中，$\overline{X}$ 是样本的平均数，μ 是总体的平均数，SE 是样本平均数的标准差，S 是样本的标准差，n 是样本数。举例来说，我们要估计北京市家庭每个月用于购买礼品的平均支出



(μ)，抽取样本，$n=600$，$\overline{X}=55$ 元，$S=11.5$ 元，则总体平均数95%置信度下的置信区间为：

$$\mu=55\pm1.96\times(12.5/\sqrt{600})=55\pm1.0$$

即：$54<\mu<56$

由此可估计出，北京市家庭平均每个月用于购买礼品的支出介于54元和56元之间，这个估计的置信度为95%，错误可能性为5%。

(2) 总体比例的区间估计

设置信度为95%，总体比例的置信区间为：

$$\mu=P\pm1.96SE=P\pm1.96\left(\sqrt{\frac{P(1-P)}{n}}\right)$$

置信度为99%，总体比例的置信区间为：

$$\mu=P\pm2.58SE=P\pm2.58\left(\sqrt{\frac{P(1-P)}{n}}\right)$$

其中 P 是样本中的比例，SE 是样本比例的标准差（抽样分布的标准差），n 是样本数。

举例来说，我们要研究某市有多少单亲家庭，从一个随机样本（$n=400$）中发现有9%的单亲家庭，即 $P=9\%=0.09$，在95%的置信度下，则总体比例 π 的置信区间为：

$$\pi=0.09\pm1.96\times\sqrt{\frac{0.09(1-0.09)}{400}}=0.09\pm0.028$$

即：$6.2\%<\pi<11.8\%$

由此推断，在该市的全部家庭中有6.2%至11.8%的家庭是单亲家庭。这个估计的置信度是95%。

(3) 使用SPSS计算总体平均数的置信区间

①利用Explore过程求总体平均数的置信区间

• 打开Analyze菜单，单击Descriptive Statistics子菜单下的

Explore 菜单项，打开 Explore 对话框；

• 该对话框左侧为变量名列表框，其中列出了当前数据文件中所有变量的变量名称；

• 在对话框右侧的 Dependent 列表框中输入待估计的变量名称，例如，输入代表北京市居民家庭平均月收入的变量；

• 单击 Statistics 按钮，打开 Explore：Statistics 对话框，选择 Descriptives 核选框，并在 Confidence Interval for Means 窗口中输入需要的置信度，例如 95%。

按以上步骤计算的北京市民家庭平均月收入 95% 置信区间的结果见表 11-10。

表 11-10　北京市民家庭平均月收入的统计量描述表

Descriptivees

			Statistic	Std. Error
家庭月收入 Mean			3644.85	481.12
	95% Confidence	Lower Bound	2696.10	
	Interval for Mean	Upper Bound	4593.60	
	5% Trimmed Mean		2587.89	
	Median		2600.00	
	Variance		46295063	
	Std. Deviation		6804.05	
	Minimum		500	
	Maximum		70000	
	Range		69500	
	Interquartile Range		1000.00	
	Skewness		7.234	.172
	Kurtosis		59.507	.342

上表中显，示样本平均家庭月收入为 3644.85 元，北京市



民总体家庭平均月收入 95% 的置信区间为［2696.10，4593.60］元。

②利用单样本 t 检验过程求总体平均数的置信区间

• 打开 Analyze 菜单，单击 Compare Means 子菜单下的 One-Sample T Test 菜单项，打开 One-Sample T Test 对话框；

• 对话框左侧为变量名列表框，其中列出了当前数据文件中所有变量的变量名称；

• 对话框的右侧 Test Variable（s）框，用于选定需要进行分析的变量。例如，将代表北京市居民家庭月收入的变量放入其中；

• Test Value 窗口，用于输入总体平均数，缺省值为 0；

• 单击 Option 按钮，在 Confidence Interval for Means 窗口中输入需要的置信度，例如 95%。

• 按照以上操作，计算结果如表 11-11 所示。

表 11-11　北京市民家庭平均月收入的单样本 t 检验成果表

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2 - tailed)	Mean Difference	Lower	Upper
家庭月收入	7.576	199	.000	3644.85	2696.10	4593.60

上表中，最后两列即为北京市民总体家庭平均月收入置信度为 95% 的置信区间上下限，即［2696.10，4593.60］元。

置信区间经常会用在精确新闻的研究中，例如，在预测美国大选的结果时，预测某位候选人的得票情况时，常常是给出

一个置信区间。预测这个候选人的得票情况在某个范围之内。

三、假设检验

假设检验的方法在统计推论上较为常用，即先有一个关于总体情况的假设，然后抽取一个随机样本，然后计算样本的统计量来检验假设。

1. 有关假设检验的基本知识

(1) 零假设和备选假设

研究一般都要先做某种假设，所谓假设就是关于总体的一个声明，即假定总体中存在某些情况，它是可以通过从该总体中抽取的一个随机样本来检验的。

零假设 H_0 通常表示为没有差异的假设，也叫原假设。例如，在精确新闻研究中，可能要确定媒体宣传是否提高了民众对于 WTO 的认知程度，为此，可设立如下零假设：

H_0：在媒体宣传前后，民众对于 WTO 的认知程度没有明显差异

即　H_0　　　　　$\square = \pi_2 - \pi_1 = 0$

其中 π_1 和 π_2 分别表示媒体宣传前后，民众中知道 WTO 的比例。

备选假设 H_1 是和零假设同时设定的。如果零假设被拒绝，那么备选假设就可能被接受。针对刚才设立的零假设，备选假设可以有以下两种方式：

H_1：在媒体宣传前后，民众对于 WTO 的认知程度有显著的差异

即　H_1：　　　　$\square = \pi_2 - \pi_1 \neq 0$

或　H_1：在媒体宣传之后，民众对于 WTO 的认知程度有



显著提高

即　H_1：　　　　　　$\square = \pi_2 - \pi_1 > 0$

上述两个备选假设是不同的，第一个没有指明变化的方向，对应的假设检验称为双侧检验；第二个隐含着所期望的变化方向，对应的检验称为单侧检验。如果在研究中，缺乏明确的理论或经验，在设立研究假设时，无法确定方向，就要用双侧检验。

检验假设的基本原则是直接检验 H_0，因而间接地检验 H_1，目的是排除抽样误差的可能性。当然，即使否定了 H_0，也不能说 H_1 一定对。

(2) 拒绝域和显著度

对于原假设 H_0 的检验，是根据抽样分布的原理。我们举例说明检验 H_0 的方法。

原假设 H_0：对某项制度同意人数（P）＝反对人数（Q）

备选假设 H_1：同意人数≠反对人数

随机抽取 10 个样本，如果 H_0 是对的，即 P＝Q，则可以计算抽样分布如下：

P（10）＝0.001　　P（9）＝0.010　　P（8）＝0.044

P（7）＝0.117　　P（6）＝0.205　　P（5）＝0.246

P（4）＝0.205　　P（3）＝0.117　　P（2）＝0.044

P（1）＝0.010　　P（0）＝0.001

其中 P（x）表示样本中有 x 个人同意的概率。可见，所抽样本中 10 个人都同意的可能性只有千分之一，全部不同意的机会也只有千分之一；9 个或 9 个以上同意的机会是：P（9）＋P（10）＝ 0.011，同样 1 个或 1 个以下不同意的可能性也是 0.011。

现在，从所抽取的样本中计算同意该制度的人数，发现是9个。按照上面的理论，如果 H_0 是对的，即 $P=Q$，那么在样本中得到9个或9个以上人同意的机会只有0.011，机会非常微小，但我们却在样本中得到了，那么 H_0 的正确性就大大可疑了。所以我们可以拒绝 H_0。

但是如果在样本中有7个人同意，从上面计算的抽样分布可见，样本中得到7个或7个以上人同意的机会是0.161。如果我们认为这个机会不算小，我们就不能从有7个人同意的样本来拒绝 H_0。

在没有对数据进行分析之前，应该先确定在什么情况下可以拒绝 H_0。例如，我们事先设定，如果所抽取样本中的同意人数是在"9个或以上"这个范围内，则拒绝 H_0。这个事先设定的范围，就叫做否定域。否定域是在抽样分布内一侧或双侧的一个区域，如果样本统计量落在此区域内，则拒绝 H_0。

与拒绝域相关的一个概念是显著度，表示样本的统计量落在拒绝域内的概率（可能性）。

在进行研究时，通常是先决定显著度的大小，在计算出样本统计量之后，检验该统计量是否落入显著度所代表的拒绝域内，是的话，就拒绝 H_0。显著度的选择，决定于研究的需要，一般选取显著度为0.05。显著度越小，越难否定 H_0。

（3）假设检验的分类

根据待分析数据和待检验假设的类型，要确定适当的假设检验的方法。假设检验基本上分为两大类：参数检验法和非参数检验法。

参数检验法要求对研究总体的性质作一些基本假定，如服从正态分布或要求方差相等。需要根据样本统计量进行推断的



总体未知量，例如，总体平均数 μ 或总体比例 π。这些总体的值被称为“参数”，因此采用的假设检验方法叫做参数检验。参数检验法一般要求变量具有定距测量层次。如果总体具备所需的条件，则参数检验就能相当准确地判别零假设的正确性。但是在使用参数检验法得到的结论包含着一些限定或条件。

非参数检验法又称为分布自由检验法，因为它不涉及总体分布情况，非参数检验法不要求变量具有定距测量层次，因而适用于推论定类或定序的数据。

后面将介绍几种常用的参数检验法，对非参数检验方法感兴趣的读者，可参考其他的书籍。

(4) 假设检验的基本步骤

①陈述零假设 H_0 和备选假设 H_1；

②选择适当的假设检验方法，根据备选假设确定检验是单侧的还是双侧的；

③确定样本量 n 和显著度的大小；

④先假定 H_0 是真的，然后根据零假设、对应的抽样分布理论、样本量 n 和显著度的大小，确定拒绝零假设的拒绝域；

⑤抽样，采集数据；

⑥计算检验统计量，如样本平均数、样本比例、样本标准差等，或计算相应的置信区间等；

⑦根据相应的公式，计算检验统计量的标准误差，从而构成一个比值作为最终检验统计量，使之与临界值比较，是否在拒绝域内，或直接计算 H_0 的概率值，用来进行检验；

⑧作出是否拒绝零假设的决定：

• 如果最终检验统计量落入拒绝域内，则拒绝零假设

• 如果最终检验统计量（的绝对值）大于临界值，则拒绝零假设



• 如果计算出的概率值小于显著度，则拒绝零假设

• 如果零假设的值落在置信区间之外，则拒绝零假设；

⑨给出结论。如果零假设没有被拒绝，那么最好将结论叙述为“无法拒绝零假设”。

使用统计软件可以方便地进行假设检验，需要注意的是精确新闻记者在写有关假设检验的结论时，要清楚阐述备选假设 H_1、显著度、检验用的是单侧还是双侧，做假设检验的结论时需要慎重。

2．*常用参数检验法*

(1) 单样本平均数的假设检验

当研究假设涉及的是总体中的一个平均数时，零假设和备选假设可以分别表示为：

$$H_0: \mu = \mu_0$$

$$H_1: \mu > \mu_0 \ (\text{或} \ \mu \neq \mu_0, \ \mu < \mu_0)$$

例如：H_0：北京市民的家庭平均月收入为 2000 元（$\mu = \mu_0 = 2000$）

H_1：北京市民的家庭平均月收入大于 2000 元（$\mu > 2000$）

①Z 检验法

在样本量比较大的情况下，样本平均数的抽样分布近似正态分布，其平均数等于 μ_0，而标准误差（SE）则等于总体的标准误差除以样本数的平方根，即：

$$SE = \sigma/\sqrt{n}$$

但是很多时候，我们无法得到总体的标准差，所以可以用样本的标准差来估计：

$$SE = S/\sqrt{n}$$

检验 H_0 的最终检验统计量是用 Z 比值（Z 检验法）：



$$Z=\frac{\overline{X}-\mu^{0}}{S/\sqrt{n}}$$

假设我们抽了一个随机样本，按照上面的公式计算 Z 比值为 2.10，通过查正态分布表可知在整个抽样分布中应该有 96.42%的样本平均数是在 ± 2.10 之间的范围内。也就是说，我们抽取一个样本，其样本平均数落在这个范围以外的可能性只有 3.58%。可能性这么小，但却在我们的样本中出现，说明零假设可能不成立，应该拒绝。也就是说，假如我们所抽取的一个随机样本的平均数在以假设 H_0 正确为基础的抽样分布中出现的可能性很小，那就应该拒绝 H_0。

这个可能性要小到什么程度才否定 H_0 呢？研究中一般采用 $p\leq0.05$，也就是显著度取为 0.05，显著度确定后，拒绝域的位置和大小取决于检验是单侧的还是双侧的。

举例说明 Z 检验法的应用，例如，考察某个学校学生数学成绩的平均数。零假设和备选假设分别是：

H_0：全校学生数学考试的平均成绩为 60 分（$\mu=60$）

H_1：全校学生数学考试的平均成绩不等于 60 分（$\mu\neq60$）

为检验假设，从全校同学中抽取一个随机样本，$n=120$，发现样本平均数为 $\overline{X}=63$，标准差 $S=19.5$，要求显著度为 0.05，由于备选假设 H_1 没有指明方向，所以需要做双侧检验。由正态分布表可查得此时拒绝域为 $|Z|\geq1.96$。

将样本平均数和标准差代入，计算 Z 比值，得出：

$$Z=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}=(63-60)/(19.5/\sqrt{120})=1.69$$

因为按照样本数据计算的 Z 比值 = 1.69，不在拒绝域内，所以不能拒绝 H_0。即不能说全校学生的数学考试平均成绩不

等于60分。

②t检验法

Z检验法要求样本量要比较大，如果样本量比较小（$n<30$），就要采用t检验法，计算t比值的公式如下：

$$t=\frac{\overline{X}-\mu_0}{S/\sqrt{n}} \qquad \text{自由度 } df=n-1$$

其中μ_0是假设的总体平均数，$\overline{X}$是样本平均数，S是样本的标准差，n是样本数。

t的抽样分布形状取决于自由度，自由度df越小，样本量越小，则t分布越扁平；自由度越大，样本量越大，t分布越高耸而且接近正态分布。自由度不同，对同一个显著度，拒绝域也会不同，不同显著度和自由度的t值，可以从t分布表中查到，也可以应用统计软件来计算。

正如前面所讲，Z检验法和t检验法的区别，在于前者基于正态抽样分布，要求样本量大，后者基于t分布，可用于小样本。但是当样本量增大时，t分布会逐渐接近正态分布，这时t检验法和Z检验法的分别就不大了。因此无论样本量大小如何，都可以使用t检验法。在精确新闻研究中，t检验法的应用更广泛一点。所以以下将着重介绍t检验法。

③使用SPSS进行单个样本平均数的t检验

使用SPSS中的单样本t检验过程（One-Sample T test）可以实现对单个样本平均数的t检验，这个过程主要检验单个样本的平均数与假设的总体平均数之间是否存在差异。

执行One-Sample T test过程，SPSS将显示每个检验变量的样本平均数、标准差和样本平均数的标准差；显示t比值、双侧检验概率值；显示样本值与假设总体值之差的95%置信区间。下面以北京市居民家庭平均月收入的研究为例，介绍



SPSS 的单样本 t 检验过程：

H_0：北京市民的家庭平均月收入为 3000 元（$\mu=\mu_0=3000$）

H_1：北京市民的家庭平均月收入不等于 3000 元（$\mu\neq3000$）

• 打开 Analyze 菜单，单击 Compare Means 子菜单下的 One-Sample T Test 菜单项，打开 One-Sample T Test 对话框；

• 对话框左侧为变量名列表框，其中列出了当前数据文件中所有变量的变量名称；

• 对话框的右侧 Test Variable (s) 框，用于选定需要进行分析的变量，将代表北京市居民家庭月收入的变量放入其中；

• Test Value 窗口，用于输入总体平均数 3000；

• 单击 Option 按钮，在 Confidence Interval for Means 窗口中输入需要的置信度，例如 95%。

• 按照以上操作，计算结果如表 11-12 所示。

表 11-12　北京市民家庭平均月收入的 t 检验结果

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
家庭月收入	200	3644.85	6804.05	481.12

(a)

One-Sample Test

	Test Value = 3000					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2 - tailed)	Mean Difference	Lower	Upper
家庭月收入	1.340	199	.182	644.85	- 303.90	1593.60

(b)

其中表 11-12（a）给出了一些样本统计量，有效样本数 $n=200$，样本平均数 $\overline{X}=3644.85$ 元，样本标准差 S 为 6804.05 元，样本平均数的标准差 $SE=481.12$ 元。

表 11-12（b）显示了对这个单样本平均数的检验结果，$t=1.340$，自由度为 199，双侧检验概率值 $p=0.182$，样本平均数和假设总体平均数之差为 644.85，平均数之差的 95% 置信区间为 [-303.90, 1593.60] 元。由双侧检验概率值 $p=0.182>0.05$ 可知，不能拒绝原假设。即不能说北京市民家庭平均月收入不等于 3000 元。

如果 SPSS 计算出的概率值小于研究设定的显著度，则可以得出拒绝原假设的结论。

（2）单样本比例的假设检验

①检验方法

如果需要检验的是总体的一个百分比（或比例），可以用 Z 检验法检验假设，方法与单样本平均数的检验类似。零假设和备选假设（双侧和单侧的）分别表示为：

H_0：北京市居民中拥有大专及以上学历的比例为 8%（$\pi=\pi_0=8\%$）

H_1：北京市居民中拥有大专及以上学历的比例不等于 8%（$\pi=\pi_1\neq 8\%$）或

H_1：北京市居民中拥有大专及以上学历的比例高于 8%（$H_0>\pi_0$）

检验公式为：$Z=(P-\pi_0)/SE=\dfrac{P-\pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}$

其中 π_0 是所假设的总体比例，P 是从样本中计算出的比例，n 是样本数，SE 是样本比例的标准误差。



例如，我们来检验 1960 年美国大选前盖洛普调查对总体选举预测的结果（51.0%，n = 1500）是否与选举实际结果没有显著的差异，已知肯尼迪的实际得票率为 50.1%。[①]

H_0：选举预测的肯尼迪得票率与选举实际结果没有显著差异，即 $\pi = \pi_0 = 50.1\%$

H_1：选举预测的肯尼迪得票率与选举实际结果有显著差异，$\pi = \pi_1 \neq 50.1\%$

取检验显著度为 0.05，样本比例 $P = 51.0\%$

$$Z = \frac{51.0 - 50.1}{\sqrt{50.1\ (100 - 50.1)\ /1500}} = 0.70$$

由于我们选取的显著度为 0.05，而在此水平下 Z 检验的拒绝域为 $Z > 1.96$，计算得出 $Z = 0.70 < 1.96$，所以在 0.05 的显著性水平下不能拒绝零假设，即选举预测的肯尼迪得票率与选举实际结果没有显著差异。

②用 SPSS 做单样本比例的假设检验

样本中的比例（P）其实就是某种意义下的样本平均数（$\bar{X}$）。例如，样本中有 60%的人表示同意某项新的政策，代表该问题的变量为 X，则对于同意的样本，令 $X = 1$；不同意则 $X = 0$，即将 X 转换为 0 - 1 变量，这样计算全部样本中变量 X 的平均数就为 0.6（60%）。既然比例是平均数的一种特殊情况，我们在进行比例检验的时候，将其视为平均数，就可以应用我们在前面介绍的平均数检验方法来做检验。所不同的是，对于需要检验的变量，在检验之前，要将其转换为 0 - 1 变量。

例如，变量 EDU 表示被调查者的学历程度

EDU = 1：初中及以下

---

① 柯惠新，丁立宏《市场调查与分析》，中国统计出版社，2000 年，第 258 页。

EDU = 2：高中、中专

EDU = 3：大专及以上

如果我们要检验学历在大专及以上人的比例，首先需要把变量 EDU 转换为 0－1 变量，另外定义一个 0－1 变量 X，令：

$X=1$：大专及以上

$X=0$：其他

这样定义的结果，变量 X 的样本平均数就等于变量 EDU 中大专及以上人的比例，因此对于变量 EDU 的样本比例检验问题就转换成为对变量 X 的样本平均数的检验问题。这样就可以应用我们前面介绍的在 SPSS 中检验样本平均数的方法来进行检验。

（3）两个独立样本平均数之差的检验

两个独立样本平均数分别为 $\overline{X_1}$ 和 $\overline{X_2}$，怎样判断它们的差究竟是由于抽样误差引起，还是因为两个总体平均数之间存在差异导致的呢？为此需要进行假设检验。零假设和备选假设分别为：

H_0：两个总体平均数没有差异（$\mu_1=\mu_2$）

H_1：两个总体平均数之间存在差异（$\mu_1\neq\mu_2$）

①t 检验法

检验统计量 t 比值的计算公式为：

$$t=\frac{\overline{X_1}-\overline{X_2}}{SE}$$

其中 $\overline{X_1}$ 和 $\overline{X_2}$ 分别为两个样本的平均数，SE 为两个样本平均数之差的标准差。

$$SE=\sqrt{\frac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}}\sqrt{\frac{n_1+n_2}{n_1n_2}}$$

$S_1{}^2$ 和 $S_2{}^2$ 分别为两个样本的样本方差，n_1 和 n_2 分别为两



个样本的样本数。

②使用 SPSS 进行两个独立样本平均数的 t 检验

在此以一个实例说明如何使用 SPSS 进行两个独立样本平均数的 t 检验。抽取 200 个样本，对某个态度进行评价，评价使用四级量表，我们要检验男性和女性对于这个问题是否存在态度上的差异。为此分别计算男性和女性在该问题上态度的平均得分，检验两个平均得分是否有统计上的差异。

H_0：男性和女性在该问题上的态度没有差异（$\mu_1 = \mu_2$）

H_1：男性和女性在该问题上的态度存在明显差异（$\mu_1 \neq \mu_2$）

• 打开 Analyze 菜单，单击 Compare Means 子菜单下的 Independent-Sample T Test 菜单项，打开 Independent-Sample T Test 对话框；

• 对话框左侧为变量名列表框，其中列出了当前数据文件中所有变量的变量名称；

• 对话框的右侧 Test Variable（s）框，用于选定需要进行分析的变量。将代表对该问题态度评价的变量放入其中；

• Grouping Variable 窗口，用于输入分组变量，此例中放入表示性别的变量，单击 Define Groups 按钮，弹出 Define Groups 定义分组对话框，选择 Use specified value，并在 Group1 和 2 后面输入变量“性别”的两个取值 1 和 2；

• 返回主窗口运行。

按照以上操作，计算结果如表 11-13 所示。

表 11-13　两个独立样本平均数的 t 检验结果

Group Statistics

	性别	N	Mean	Std. Deviation	Std. Error Mean
态度得分	男	68	2.26	1.06	.13
	女	132	1.84	.85	7.36E-02

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
态度得分	Equal variances assumed	10.547	.001	3.074	198	.002
	Equal variances not assumed			2.862	111.997	.005

从上面 t 检验的结果可以看出，SPSS 在进行两个独立样本平均数的 t 检验时，会给出分组统计表，包括有效样本数、平均数、样本标准差和样本平均数标准差。表中显示男性在该问题上的平均得分是 2.26；女性是 1.84。

在 t 检验表中，给出了方差相等和方差不等两种情况下的方差齐性检验结果和等均值的 t 检验结果。如果两个样本的方差相等或不等，应选择对应行的统计量作为 t 检验的结果。如果方差齐性检验的结果即表中的第三列 Sig. >0.05，则在 0.05 显著度下，方差相等的条件满足；否则即表示两个样本的方差不相等。t 检验的结果包括 t 值、自由度、双侧概值等。在表 11-13 中，方差齐性检验显著性 =0.001 <0.05，所以方差不相等，应该看第二行的 t 检验结果，$t=2.862$，双侧概值 =0.005 < 0.05，所以可以得出结论，在 0.05 的显著度下，男性和女性在该问题上态度平均得分存在显著差异，即可以拒绝原假设。

我们很容易可以把 t 比值的应用推广到其他情况的假设检验，如，两个随机样本平均数或比例之差的假设检验。我们可



以将 t 比值写成一个一般形式如下：

t =（样本估计值 - 零假设的值）/样本估计值的标准差

例如，进行对于样本平均数的检验，t 比值分子中样本估计值即为通过样本计算的样本平均数（$\overline{X}$），而零假设的值就是假设的总体平均数，分母为样本平均数的标准误差 SE。根据上式构造 t 比值，就可以对总体参数进行 t 检验了。

前而我们以单样本平均数和比例的假设检验为例，说明了假设检验的逻辑和步骤。由于精确新闻研究中，对于单一统计量或统计量之间差异的检验关心的不多，最关心的是研究两个变量之间的关系。

第四节　常用双变量分析方法

前面我们一直在讨论对单个变量的分析，而在精确新闻报道研究中，常常需要分析两个变量之间的关系。例如，研究不同受教育程度的人对于《新婚姻法》的认知和支持程度是否存在差异，男女在消费观念上是否不同，等等，对于这些问题的探讨都涉及了两个变量，因此需要进行双变量分析，研究一个变量与另一个变量之间的关系问题。在统计学中，常用的双变量分析方法有：卡方检验和交互分析；单因素方差分析；相关分析；回归分析等。

一、卡方检验和交互分析

1. 交互分析

（1）交互分析方法

所谓交互分析，就是同时依据两个变量的值，将所研究的个案进行分类。交互分析是最简单的双变量分析方法。交互分

析适用于分析两个定类变量之间的关系，当然如果降低级别处理的话，也可以用交互分析法分析较高级别的变量。

例如，某媒体希望研究北京市青年人的最大志愿是什么，为此在北京市抽取了200名青年样本，调查他们的最大志愿，表11-14是样本频数表，由此可见，样本中34.5%的青年人的最大志愿是拥有“幸福家庭”；56.5%的人说最大志愿是拥有“成功事业”；9%的人认为最大志愿是要有“丰富经历”。

表11-14 青年人的最大志愿频数表

青年人的最大志愿

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	幸福家庭	69	34.5	34.5	34.5
	成功事业	113	56.5	56.5	91.0
	丰富经历	18	9.0	9.0	100.0
	Total	200	100.0	100.0	

频数表告诉我们一些结果，但是如果我们想研究青年人的最大志愿与其他变量的关系，比如，与受教育程度的关系，就可以采用交互分析。

（2）使用SPSS进行交互分析

• 打开Analyze菜单，单击Descriptive Statistics子菜单下的Crosstabs菜单项，打开Crosstabs对话框；

• Row（s）：行变量栏，对于上面这个问题，将代表青年人最大志愿的变量放入该栏；

• Column（s）：列变量栏，将代表受教育程度的变量放入该栏；

• 单击OK键继续。

在SPSS中按照以上步骤计算上例中两个变量的交互分析结果如表11-15所示：



表 11-15　青年人的受教育水平与最大志愿

青年人的最大志愿 * 受教育程度 Crosstabulation

Count

		受教育程度			Total
		低	中	高	
青年人的最大志愿	幸福家庭	24	19	26	69
	成功事业	24	32	57	113
	丰富经历	3	9	6	18
Total		51	60	89	200

表 11-15 即是利用 SPSS 将 200 名样本按照其受教育水平和最大志愿交互分类的结果，在表 11-15 中，可以清楚地知道在不同受教育水平下，青年人最大志愿的频数分布情况。例如，在样本中，受教育水平低的有 51 人，其中 24 人的最大志愿是“幸福家庭”。这种表称为交互表，又称列联表。交互表的大小 = r * c，其中 r 表示表的行变量取值的数目，而 c 表示列变量取值的数目。表 11-15 是一个大小为 3 * 3 的交互表，其中行变量为最大志愿，列变量为受教育水平。

但是表 11-15 有一个问题，就是没有办法比较不同情况下的频数分布，例如，我们看到，在受教育水平高的青年中有 26 个人的最大志愿是“幸福家庭”，而中等教育水平中有 19 人的最大志愿是“幸福家庭”，前者多于后者，是否表示受教育水平高的人更多地以拥有“幸福家庭”为最大志愿呢？表 11-15 的最后一行显示，受教育水平高的样本基数为 89 人，也多于受教育水平中等的 60 人，两者的基数不同，所以我们不能轻易下结论。为了在相同的水平上进行比较，就必须将基数标准化，方法就是将频数转换为百分比。

在 SPSS 中进行交互分析时，可以同时给出频数对应的百分比。在 Crosstabs 对话框中，单击 Cells 按钮，打开 Cell Dis-

play 对话框，在 Pereentages 中选择 Row 和 Column 两个复选框，其余采用默认值，重新计算表 11-15 的数据，结果见表 11-16：

表 11-16　青年人的受教育水平与人生志愿

青年人的最大志愿 * 受教育程度 Crosstabulation

			受教育程度			Total
			低	中	高	
青年人的最大志愿	幸福家庭	Count	24	19	26	69
		% within 青年人的最大志愿	34.8%	27.5%	37.7%	100.0%
		% within 受教育程度	47.1%	31.7%	29.2%	34.5%
	成功事业	Count	24	32	57	113
		% within 青年人的最大志愿	21.2%	28.3%	50.4%	100.0%
		% within 受教育程度	47.1%	53.3%	64.0%	56.5%
	丰富经历	Count	3	9	6	18
		% within 青年人的最大志愿	16.7%	50.0%	33.3%	100.0%
		% within 受教育程度	5.9%	15.0%	6.7%	9.0%
Total		Count	51	60	89	200
		% within 青年人的最大志愿	25.5%	30.0%	44.5%	100.0%
		% within 受教育程度	100.0%	100.0%	100 0%	100.0%

先解释一下这张表中的数据，表中每个单元格里有 3 个数字，其中第一个数表示频数，意义同表 11-15，例如，从表中可见，有 24 个受教育水平比较低的人表示自已的最大志愿是拥有“幸福家庭”，行合计给出 69 个人最大志愿是“幸福家庭”，113 个人是“成功事业”，18 个人是“丰富经历”；列合计给出各种受教育水平的人数，例如，受教育程度高的有 89 人，等等。总的样本含量为 $n=200$ 人。

单元格中第二个为行百分比，是频数相对于行合计计算的百分比。例如，最大志愿是拥有“幸福家庭”的青年人中，受教育水平比较低的占 34.8%，受教育水平中等的为 27.5%，受教育水平高的为 37.7%，每 -行的行百分比合计为 100%。

单元格中第三个为列百分比，是频数相对于列合计计算的



百分比。例如，在受教育水平比较低的年轻人中，最大志愿是“幸福家庭”的比例占 47.1%，是“成功事业”的占 47.1%，等等。每一列的列百分比合计为 100%。

那么到底按照哪一个百分比进行比较呢，常用的规则是根据自变量的方向，因为研究的目的是了解自变量对因变量的影响，所以要计算在自变量取不同值情况下，因变量的变化。表 11-16 是研究受教育水平对人生志愿的影响，所以应该看不同教育水平中，人生志愿的变化情况。也就是说，在这个例子中，列百分比是我们所关注的。当然这个规则不是绝对的，应用时要根据实际情况具体分析。

从表 11-16 的列百分比中可以看到，受教育水平不同的人，他们的人生志愿有所差别，特别是受教育水平越高，以“幸福家庭”为人生志愿的比例越低，而认为“成功事业”是自己最大志愿的比例越高。

交互分析虽然可以用来探讨两个变量之间的关系，但它呈现的只是各个自变量和因变量在数值上的差异，而没有显示这些差异在统计上是否达到显著水平。也就是说，从表 11-16 中看到的不同受教育水平青年人之间的差异，到底是真正的差异，还是由于抽样误差造成的。为此应该进一步采取其他更为有力的统计分析方法，以显示变量在数值上的差异是否在统计上也是显著的，并对此进行假设检验。

2. 卡方检验

(1) 卡方检验

常用于交互分析中的检验方法是卡方检验。卡方检验是非参数检验中的一种，应用时不涉及总体分布的类型。卡方检验可以用来检验变量间的独立性（是否相关）检验，适用于两个变量都是定类变量的情况。

在进行卡方检验前，研究者需提出研究假设。其中零假设通常假设两个变量间相互独立，没有关联性，而备选假设则认为两个变量不独立。在交互分析中使用卡方检验可以显示变量间的关系或差异，在统计上是否显著。

卡方统计量的计算公式为：

$$\chi^2 = \sum\sum \frac{(f_0 - f_e)^2}{f_e}$$

其中 f_0 为观测到的频数，f_e 为零假设 H_0 为真时的期望频数，两个累加号表示将交互表中的数按行、列全部相加。卡方统计量的自由度为 $df = (c-1)(r-1)$。r、c 分别表示交互表的行数和列数。

(2) 使用 SPSS 进行交互分析的卡方检验

我们以前面关于青年人教育水平和人生志愿关系研究为例，介绍用 SPSS 进行卡方检验的过程。

在这个例子中，零假设和备选假设分别为：

H_0：受教育程度与人生志愿相互独立

H_1：受教育程度不同的人在人生志愿上存在显著差异

在 SPSS 中进行交互分析的卡方检验，只需在 Crosstabs 对话框中，单击 Statistics 按钮，打开 Statistics 对话框，选择 Chi-square 和 Contingency coefficient 复选框，SPSS 在进行交互分析的同时，就会进行卡方检验。

针对上例的卡方检验结果见表 11-17：

表 11-17　卡方检验

Chi-Square Tests

	Value	df	Asymp. Sig. (2 – sided)
Pearson Chi-Square	8.455ᵃ	4	.076

Likelihood Ratio	7.993	4	.092
Linear-by-Linear Association	2.315	1	.128
N of Valid Cases	200		

a. 1 cells (11.1%) have expected count less than 5.

The minimum expected count is 4.59.

表11-17显示卡方值 $\chi^2 = 8.455$，这个值需要和最后一列给出的显著度一起分析才有意义。卡方值越大，显著度就会越小。如果研究要求的检验显著度为0.05，则Sig.值小于0.05时，就表示被检验的两个变量之间存在显著的关联。而在这个例子中，$p = 0.076 > 0.05$，所以不能拒绝原假设，即受教育水平和人生志愿之间是相互独立的。表11-17中还给出自由度 $df = 4$，卡方检验表下面的注释说明有一个单元格（11.1%）的期望值小于5，最小的期望值是4.59。如果有20%的期望值小于5，则不适于使用卡方分析。可能需要研究者进行重新分类。

（3）注意事项

卡方检验虽然简单易用，但也有其自身的局限性。

首先卡方值（χ^2）随分类的不同而改变。由于交互分析是将数据按照类别整理，分类的不同可以改变卡方值，甚至可能导致相反的结论。例如，上例中的受教育水平，可以分为高、中、低三类，也可以分为高、低两类。因此在分类时，最后有非统计理论支持。

其次在做卡方检验时，样本量不能太小，也不宜过大。在精确新闻调查中，样本含量一般都会超过50，这没有问题；

但样本量过大，有时得到的结果将失去意义，样本量越大，越容易得到拒绝原假设 H_0 的结果，为了消除样本含量的影响，一般在进行卡方检验时，还要参考列联系数。

另外列联表中期望频数小于5的个数不能太多，建议所有期望频数不小于5，放宽要求也不能超过20%。

由于卡方检验有以上的一些局限性，因此在使用卡方检验解释数据时，一定要谨慎。

二、单因素方差分析

前面提到使用 t 检验的方法可以检验两个总体的平均数是否相等，而我们在研究中涉及的自变量经常会有3个或3个以上的类别。比如，前面我们研究青年人的受教育水平和人生志愿的关系，受教育水平就分为3类：高、中、低。如果需要比较这些类之间平均数的差异时，例如，比较不同受教育水平的人对某事物的评分是否有差异等等，就需要采用方差分析的方法。方差分析一般要假定所比较的总体具有相同的方差和正态分布，不过方差分析在更宽的条件下也还是近似有效的。另外相关分析要求是随机样本，分析中有一个变量是定距变量。

1. 方差分析的基本方法

一般而言，进行方差分析时，零假设和备选假设分别为：

H_0：各个总体的平均数没有差异（$\mu_1 = \mu_2 = \cdots = \mu_n$）

H_1：各个总体的平均数之间不完全相同

使用方差分析方法进行假设检验要做 F 检验。我们以一个例子说明 F 比值的计算方法。

例如，我们在不同受教育程度的人中间各随机抽取了8个样本，调查他们看电视的时间（小时），数据如下：



样本 教育	1	2	3	4	5	6	7	8	合计	平均数
高	2	4	5	6	3	5	3	8	36	4.5
中	9	5	8	12	14	3	8	13	72	9
低	3	5	7	8	11	10	5	3	52	6.5

从中可见，受教育水平不同的样本在看电视时间上的平均值是不同的，但是这种不同是由于抽样误差造成的，还是因为不同受教育水平的总体之间存在差异呢？为此需要根据样本数据进行假设检验。在这个例子中，零假设和备选假设分别为：

H_0：受教育水平不同的人在看电视的平均时间上没有差异（$\mu_1=\mu_2=\mu_3$）

H_1：受教育水平不同的人在看电视的平均时间不完全相同

要计算 F 比值，需要计算两种变差，一种是组间变差，一种是组内变差。

(1) 组间变差

组间变差主要是估计各组平均数之间的差异，为此要先求出 3 个样本的总平均数 $\bar{X}$，

$$\bar{X} = 1/c\sum \bar{X}_i = 1/3(4.5+9+6.5) = 6.67$$

其中 c 表示分组数，即自变量的取值数。然后计算各组样本平均数 $\bar{X}_i$ 相对于总平均数 $\bar{X}$ 的方差

$$S_x{}^2 = \frac{1}{c-1}\sum(\bar{X}_i-\bar{X})^2$$

$$= 1/(3-1)\times[(4.5-6.67)^2+(9-6.67)^2+(6.5-6.67)^2] = 5.08$$

S_x^2 描述了各组平均数和总平均数的差异，组间变差 BSS $= nS_x^2 2\ (c-1)$，n 为每组内的样本数，BSS 它代表了由于各组平均数之间的差异造成的变差。

（2）组内变差

组内变差测量每组样本内观测值偏离其各组平均数的程度。定义联合方差 S_p^2，

$$S_p^2 = \frac{\sum(X_1 - \overline{X_1})^2 + \cdots + \sum(X_c - \overline{X_c})^2}{c(n-1)}$$

$$= [(2-4.5)^2 + \cdots + (8-4.5)^2 + (9-9)^2 + \cdots + (13-9)^2 + (3-6.5)^2 + \cdots + (3-6.5)^2]/21$$

$$= 9.23$$

组内变差 WSS $= S_p^2 c\ (n-1)$，代表由于随机波动造成的变差。

（3）计算 F 比值

$F = nS_x^2 / S_p^2 = 5.08 \times 8/9.23 = 4.4$

其中分子的自由度 $df_1 = c-1 = 2$

分母自由度 $df_2 = c\ (n-1)\ = 21$

（4）检验零假设

在 0.05 显著度下，查 F 分布表，可知此时的拒绝域为 $F > 3.47$，前面计算出 $F = 4.4 > 3.47$，所以可以得出结论：拒绝零假设 H_0，即不同受教育水平的人看电视的平均时间是不完全相同的。

以上即是利用 F 检验进行方差分析的基本步骤。

如果用 TSS 代表每一个数据对总平均数 $\overline{X}$ 的总变差，可以证明：

TSS = BSS + WSS



即总变差＝组间变差＋组内变差

总自由度＝分子自由度＋分母自由度

而 $F=\frac{BSS/df_1}{WSS/df_2}=\frac{\text{组间方差}}{\text{组内方差}}=\frac{\text{总方差中可以解释的部分}}{\text{总方差中不能解释的部分}}$

用组间变差和组内变差分别除以各自的自由度，就得到组间方差和组内方差，根据各组可能属于不同总体可以解释组间的方差，而组内方差是由于随机波动造成的，无法系统解释。所以我们看到，所谓方差分析实际上是将全部方差分解为两个部分：可解释的方差和不能解释的方差，而 F 比值就是可解释的方差和不能解释的方差之比。

2．使用 SPSS 进行方差分析

同样以一个例子说明如何使用 SPSS 进行方差分析。在某市随机抽取了 200 名样本，研究他们对实行夏时制的态度，评价采用四级量表，其中 1 表示"非常同意"，2 表示"有点同意"，3 表示"有点不同意"，4 表示"非常不同意"。在进行数据分析时，将表示同意程度的变量看做定距变量，计算不同人群的平均得分。现在使用方差分析的方法考察不同受教育水平人在此问题上的态度是否有差异。

H_0：不同受教育水平的人对于实行夏时制的同意程度没有差异（$\mu_1=\mu_2=\mu_3$）

H_1：不同受教育水平的人对于实行夏时制的同意程度不完全相同

• 打开 Analyze 菜单，单击 Compare Means 子菜单下的 One-Way ANOVA 菜单项，打开 One-Way ANOVA 对话框；

• 对话框左侧为变量名列表框，其中列出了当前数据文件中所有变量的变量名称；

• 对话框的右侧 Dependent List 框，用于选定需要进行分

析的因变量。将代表对该问题态度评价的变量放入其中；

• Factor 窗口，用于输入分组变量，此例中放入表示受教育水平的变量，该变量将人们的教育水平分为3类：高、中、低。

• 单击 Options 按钮，打开 Options 对话框，该对话框用于选择需要输出的统计量。比如，Descriptive 复选项，输出描述性统计量，如：有效样本数、样本平均数、标准差、平均数的标准误差、最大值、最小值、各组中每个因变量的95%置信区间；Himogeneity-of-variance 复选项用于对方差进行齐次性检验；

• 返回主窗口运行。

使用 SPSS 进行单因素方差分析的结果如表 11-18 所示：

表 11-18　单因素方差分析结果

Descriptives

Q8J　对实行夏时制的同意程度

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1 低	86	1.84	.89	1.E-01	1.65	2.03	1	4
2 中	61	1.89	.86	.11	1.67	2.11	1	4
3 高	51	2.37	1.04	.15	2.08	2.66	1	4
Total	198	1.99	.95	7.E-02	1.86	2.12	1	4

(a)

Test of Homogeneity of Variances

Q8J　对实行夏时制的同意程度

Levene Statistic	df1	df2	Sig.
2.830	2	195	.061

(b)



ANOVA

Q8J　对实行夏时制的同意程度

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	10.141	2	5.070	5.962	.003
Within Groups	165.839	195	.850		
Total	175.980	197			

(c)

表 11-18（a）为选择了 Options 对话框中的 Descriptive 项所得的输出结果，可见，参与分析的有效样本数为 198，三组的平均数分别为 1.84，1.89，2.37。

表 11-8（b）为选择了 Options 对话框中的方差齐次检验项所得的输出结果。

表 11-18（c）显示了单因素方差分析的结果，可见总方差来源于两个部分，即组间（Between Groups）和组内（Within Groups），其自由度分别为 2 和 195，总自由度为 197；组间变差为 10.141，组内变差为 165.839，总变差 175.980；组间方差为 5.070，组内方差为 0.850；F 值为 5.962，$p = 0.003 < 0.05$。那么可以得出结论，在 0.05 的显著度下，可以拒绝零假设，即认为不同受教育水平的人对于实行夏时制的同意程度不完全相同。

三、相关分析

前面介绍的卡方分析可以显示两个变量是否具有关联性，而 t 检验和方差分析，可以比较两个或两个以上平均数的差异。但是两个变量间关联的强度，却没有办法从这三种分析中知道，而相关分析却可以为两个变量关联的紧密程度提供一个

度量。相关分析是广为应用的统计分析方法，它计算简单，意义直观。例如，要研究中央电视台《焦点访谈》的内容与其收视率的关联程度，就可以使用相关分析。

1. 相关关系的种类

根据变量的表现形式和变化方向的不同，可以将相关关系分为：

(1) 线性相关和非线性相关

如果在相关分析中，两个变量的相关关系可以近似地描绘成一条直线，就称为线性相关。例如，如果彩票的中奖率越高，购买的人数也就越多，如果这种相关关系可以近似描绘成一条直线，就是线性相关；如果两个变量之间的关系描绘的不是一条直线，而是曲线（如抛物线、双曲线）等，就是非线性相关。例如，在百货商店门口观察人流量，如果把人流量和时间的关系表现在坐标上，就是一种曲线关系。

(2) 正相关和负相关

如果在相关分析中，自变量（X）和因变量（Y）的关系是呈正比例变化的，就是正相关。例如，上面所举彩票中奖率（X）和购买人数（Y）的关系；如果自变量（X）和因变量（Y）的关系是呈反比例变化的，就是负相关。例如，收入越高（X）的家庭，食品消费的支出越低（Y）。

2. 简单（积矩）相关

在相关分析中比较常用的是简单积矩相关。简单积矩相关系数又称为皮尔逊（Pearson）相关系数，主要是测量两个定距变量之间联系的紧密程度如何，有时，研究者也会将定序变量看做定距变量，计算其皮尔逊相关系数。

(1) 样本相关系数 r 的计算公式

样本相关系数 r 可依据下式计算：



$$r=\frac{\sum(X-\overline{X})(Y-\overline{Y})}{\sqrt{\sum(X-\overline{X})^2}\sqrt{\sum(Y-\overline{Y})^2}}$$

其中 X 为自变量的取值，$\overline{X}$ 为自变量取值的平均数；Y 为因变量的取值，$\overline{Y}$ 为因变量取值的平均数。

(2) 样本相关系数 r 的直观意义

样本相关系数 r 测量的是两个变量间线性关系的紧密程度，取值在 -1 到 +1 之间，即 $-1\leq r\leq 1$。如果相关系数等于 +1，表示两个变量呈完全的正相关［如图 11-3 (2)］；如果两个变量的相关系数等于 -1，则表示两个变量呈完全的负相关［如图 11-3 (4)］；图 11-3 (1) 和图 11-3 (3) 表示了通常的正相关和负相关，r 越接近于 +1，说明正相关关系越强；r 越接近于 -1，说明负相关关系越强；如果相关系数等于 0，那么有两种可能性，其一，两个变量之间没有任何联系［如图 11-3 (5)］，其二，两个变量间有很强的关系，但是非线性关系［如图 11-3 (6)］。因此，相关系数等于 0，并不意味着两个变量之间没有关系，而只意味着两个变量间没有线性关系，样本相关系数 r 只是线性关系的一种度量。

(3) 总体相关系数的假设检验

求出了样本相关系数 r 之后，还需要利用它来对总体的相关情况做统计推断。设总体中两个变量的相关系数用 ρ 表示，对总体相关系数进行假设检验的零假设和备选假设分别为：

H_0：总体中两个变量没有线性相关关系 ($\rho=0$)

H_1：总体中两个变量线性相关 ($\rho\neq 0$)

上述假设是采用双侧检验，如果我们有理论足以支持假设的方向，可以采用单侧检验如下：

H_0：总体中两个变量没有线性相关关系 ($\rho=0$)

H_1：总体中两个变量呈线性正（负）相关（$\rho>0$ 或 $\rho<0$）

对于总体相关系数的检验可以采用 t 检验方法，t 比值定义为：

$$t=\frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

其中 r 为样本相关系数，n 为样本数，自由度为 $n-2$。

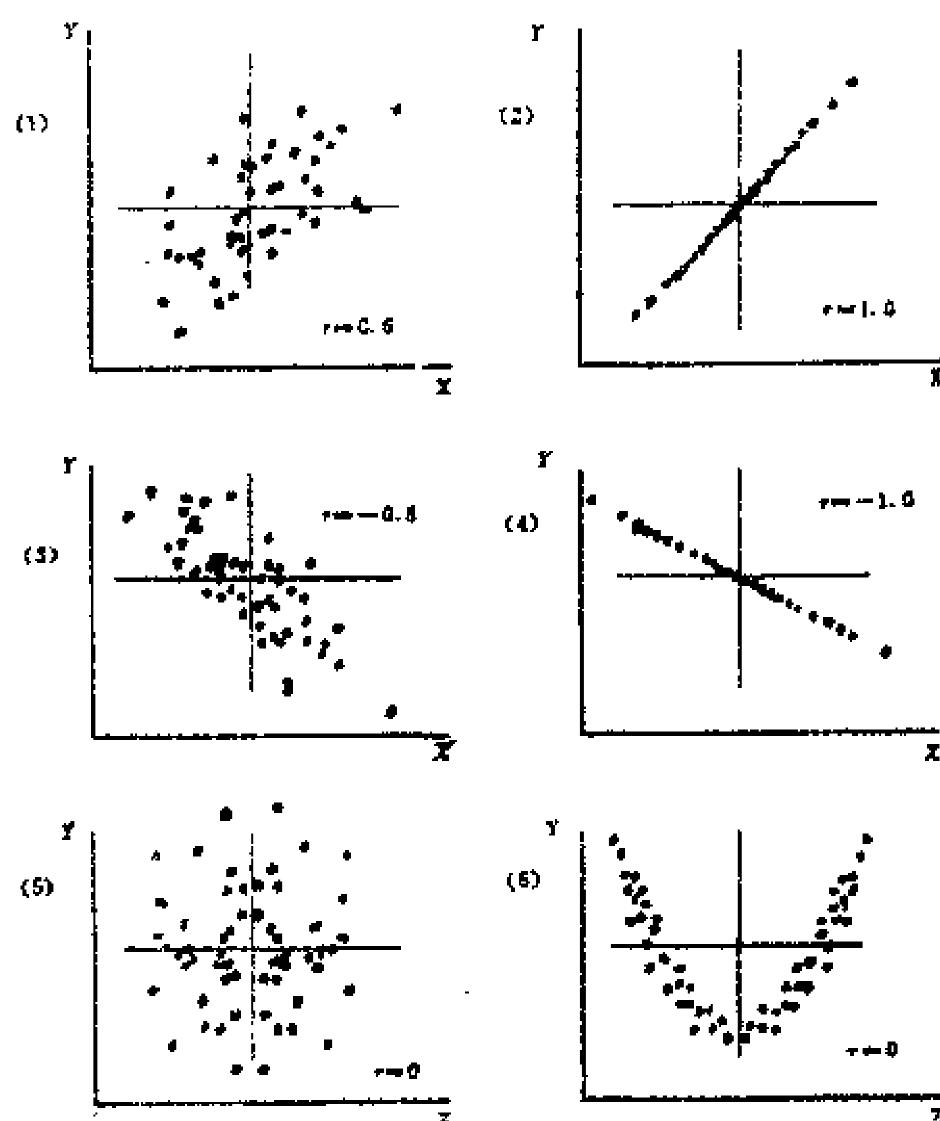


图 11-3 各种相关



3．利用 SPSS 进行相关分析

利用 SPSS 可以方便地进行相关分析，下面以一个例子来进行介绍。

例如，研究居民非商品支出与居民文化生活支出的关系，在某小区共随机抽取了 18 个样本，样本数据如下：

样本编号	月均非商品支出（元）	月均文化生活支出
1	109	87
2	104	84
3	92	72
4	102	82
5	92	73
6	113	89
7	96	74
8	99	80
9	97	76
10	105	81
11	101	82
12	108	85
13	105	83
14	99	78
15	100	80
16	94	75
17	97	80
18	95	73

前面我们看到散点图（图 11-3）对于两个变量的相关关系是一种非常直观的表示方法，我们可以先通过散点图观察居民非商品支出和文化生活支出之间的关系。具体做法是：

• 单击 Graphs 菜单中的 Scatter 菜单项，打开 Scatterplot 对

话框进行定义；

• 选择简单散点图（Simple），单击 Define 按钮，打开 Simple Scatterplot 定义对话框；

• 选择居民文化生活支出为 X 轴，居民非商品支出为 Y 轴。

按照以上步骤运行，SPSS 输出的散点图见图 11-4 所示。

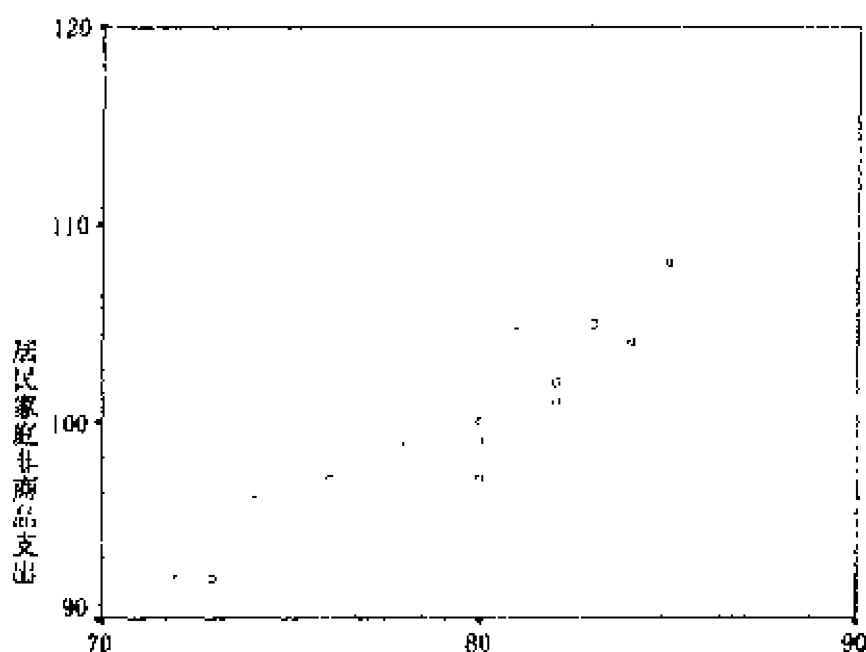


图 11-4　居民非商品支出和文化生活支出的散点图

由上图可见，居民家庭非商品支出和文化生活支出之间是有很强的线性关系。当然我们也可以不做散点图，而直接利用 SPSS 相关分析功能菜单进行相关分析：

• 打开 Analyze 菜单，单击 correlate 子菜单下的 Bivariate 菜单项，打开 Bivariate Correlations 对话框；

• 对话框左侧为变量名列表框，其中列出了当前数据文件中所有变量的变量名称；

• 对话框的右侧 Variables 框，用于选定需要进行相关分析的变量。将代表居民月非商品支出和月文化生活支出的变量放入其中；



• 在 Correlation Coefficient 下的 3 种相关系数复选框中选择 Pearson 相关系数；

• 在事先不知道相关方向时，选择 Two-tailed 双侧检验，否则选择单侧（One-tailed）检验；

• 单击 Options 按钮，可打开 Options 对话框，该对话框用于选择需要输出的统计量。比如，样本平均数、标准差和选择缺失值处理方法。

采用 SPSS 系统默认设置输出的两个变量相关分析结果见表 11-19。

表 11-19　居民月非商品支出和月文化生活支出相关分析结果

Correlations

		居民家庭非商品支出	居民家庭文化生活支出
居民家庭非商品支出	Pearson Correlation	1.000	.955 *
	Sig.（2 – tailed）	.	.000
	N	18	18
居民家庭文化生活支出	Pearson Correlation	.955 * *	1.000
	Sig.（2 – tailed）	.000	.
	N	18	18

* * . Correlation is significant at the 0.01 level（2-tailed）.

上表每个单元格中第一行显示了两个变量的皮尔逊相关系数，相关系数右上角有一个 * 符号，代表两个变量在 0.05 显著度下统计上是显著相关的，即可以拒绝零假设，两个 * 符号代表两个变量在 0.01 显著度下统计上是显著相关的，即可以拒绝零假设；第二行是对于相关系数的双侧检验概率值，第三行代表参与分析的有效样本数。由此可知居民月非商品支出和月文化生活支出的相关系数为 0.955，且在 0.01 显著度下统计

上是显著相关的。

使用 SPSS 不仅可以进行两个变量间的相关分析，而且可以同时分析多个变量两两之间的相关关系，形成含多个变量的相关表。见表 11-20：

表 11-20　多个变量相关表

Correlations

		Q1	Q2	Q4	Q6	Q7
Q1	Pearson Correlation	1.000	.048	-.036	-.021	.078
	Sig.（2-tailed）	.	.237	.373	.604	.057
	N	600	600	600	600	600
Q2	Pearson Correlation	.048	1.000	.074	.056	.133**
	Sig.（2-tailed）	.237	.	.071	.167	.001
	N	600	600	600	600	600
Q4	Pearson Correlation	-.036	.074	1.000	.064	.105**
	Sig.（2-tailed）	.373	.071	.	.117	.010
	N	600	600	600	600	600
Q6	Pearson Correlation	-.021	.056	.064	1.000	.103*
	Sig.（2-tailed）	.604	.167	.117	.	.012
	N	600	600	600	600	600
Q7	Pearson Correlation	.078	.133**	.105**	.103*	1.000
	Sig.（2-tailed）	.057	.001	.010	.012	.
	N	600	600	600	600	600

**. Correlation is significant at the 0.01 level（2-tailed）.

*. Correlation is significant at the 0.05 level（2-tailed）.

对表 11-20 的解释和表 11-19 是类似的。相关表是以对角线为中心对称的，因为 Q2 和 Q7 之间的相关系数（0.133）与 Q7 和 Q2 之间的相关系数是完全一样的。



我们看到，SPSS提供3种相关系数的计算，除了我们介绍的皮尔逊相关系数之外，还有Kendall's tau-b和Spearman相关系数。其中Kendall's tau-b用于测量两个具有对称关系的定序变量间的相关，同样适用于定序变量和定序变量间的相关。这两种相关系数的详细资料，可参看有关的统计书籍。

四、回归分析

相关分析可以使我们了解两个变量之间相互联系的紧密程度，如果两个变量相关，其中一个变量的变化，将影响另一个变量。但是从相关分析的结果，看不出这两个变量之间是如何联系的，一个变量变化，对另一个变量的影响有多大，或者说，一个变量变化，另一个变量如何跟着变化。这些问题需要通过回归分析来解决。

回归分析主要用于考察一个变量是如何和其他变量相联系的，在这里我们只讨论一个变量和另一个变量的联系。

1. 简单线性回归

我们研究两个变量 X 和 Y 之间的关系，如果变量 X 和 Y 相关，那么 X 的变化，必然会影响到 Y 也随之发生变化。例如，家庭收入 X 和家庭的食品支出 Y 这样两个变量，可以想见，当 X 发生变化时，Y 也会跟着变化，如果我们能用数学式子将这种关系表现出来，那么我们就可以根据 X 的值去预测 Y 的值，即通过一个家庭的收入，预测该家庭用于购买食品的支出大概是多少。由于家庭食品支出（Y）取决于家庭的收入（X），所以称食品支出（Y）为因变量或响应，而家庭收入（X）不依赖于家庭食品支出（Y），所以 X 是独立的，称为自变量，也叫因子或回归子。回归分析和相关分析一样，是精确新闻报道常用的分析方法。例如，在1997年10月10日《南

方周末》上刊登的"您最担心什么"的精确新闻报道中，相关分析显示，收入越高的夫妻对"配偶是否忠诚"的担忧比例越高，而对对方"工作、事业"担忧的比例越低。通过回归分析，就能够准确测度出不同收入层次（X）与配偶是否忠诚的担忧比例（Y）之间一个的数值关系。[①]

由此看来，所谓回归分析就是用样本数据拟和一个数学方程式，然后根据这个方程式，可以用自变量（X）的数值来预测因变量（Y）的数值。这里我们只研究其中最简单的一种情况，即用一条直线方程将 Y 和 X 相联系的情况表现出来，称为 Y 对 X 的简单线性回归。进行回归分析时，自变量和因变量都需要是定距或定比变量。

例如，研究者希望用女青年的受教育年限来预测或估计她们的收入。为此，调查了9名女青年，收集数据如下：

样本编号	受教育年限（年）	月收入（元）	$(X-\overline{X})$	$(Y-\overline{Y})$	$(X-\overline{X})*(Y-\overline{Y})$	$(X-\overline{X})^2$
1	2	300	-7.4	-767	5675.8	54.76
2	3	400	-6.4	-667	4268.8	40.96
3	6	500	-3.4	-567	1927.8	11.56
4	9	800	-0.4	-267	106.8	0.16
5	9	1000	-0.4	-67	26.8	0.16
6	12	1200	2.6	133	345.8	6.76
7	12	1500	2.6	433	1125.8	6.76
8	16	1800	6.6	733	4837.8	43.56
9	16	2000	6.6	933	6157.8	43.56
	$\overline{X}=9.4$	$\overline{Y}=1067$			$\sum=24473.2$	$\sum=208.24$

① 姜秀珍《新闻统计学》，新华出版社，1998年，第317页。

以上数据的散点图见图11-5所示。

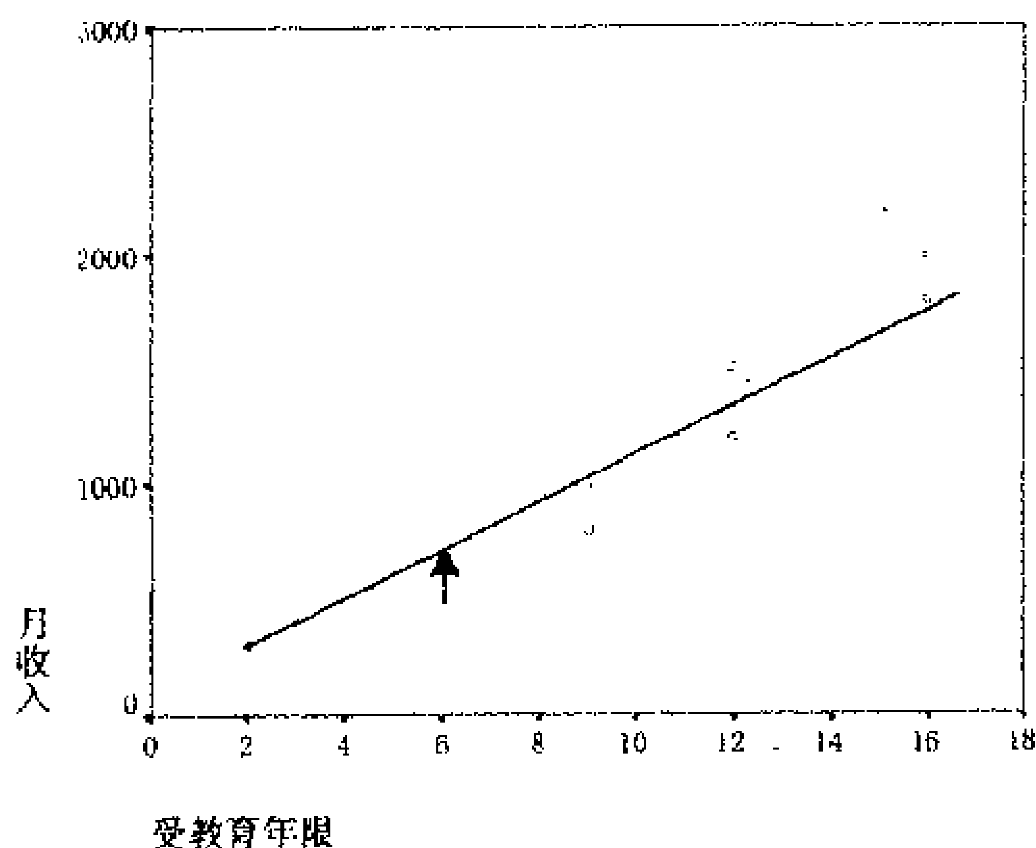


图11-5　女青年受教育年限和收入散点图

从图11-5可见，女青年的受教育年限和她们的收入之间是一种正相关的关系，我们可以用一条直线来拟和这种关系，图中画出了一条拟和直线。假设这条直线的方程是：

$\hat{Y} = a + bX$

这条拟和直线不能正好通过所有的观测样本点，因此如果用这条直线来预测的话，是有预测误差的。例如，从这条直线来看，当 $X = 6$ 时，预测值 $\hat{Y} \approx 700$ 元，而实际观测值 $Y = 500$ 元。用 d 表示预测值和观测值之间的偏差，则 $d = Y - \hat{Y}$。

2. 用最小二乘法拟和回归直线

拟和的直线可以有很多条，上图中只是画出了其中一条，那么哪一条才是最好的呢？我们当然希望拟和的直线与实际观

测值之间的偏差最小，怎样才能做到这一点呢？

(1) 最小二乘准则

拟和的直线方程形式为：$\hat{Y} = a + bX$

我们要找到计算 a（截距）和 b（斜率）的公式。依照上面提到的准则，使所有观测值和预测值的偏差 d 最小，原则上，将全部样本个案的观测值和预测值的偏差 d 相加就是总偏差，但这样做会引起正负值相抵消的问题，所以我们改为把 d 的取平方后相加。因此，如果拟和的回归直线在坐标图中的位置，能够使得是 $\sum d^2$ 最小的，就应该是最佳拟和直线了。这就是所谓最小二乘（平方）准则，即使得由回归方程和观测值计算的 $\sum d^2$ 是最小的。

$$\sum d^2 = \sum (Y - \hat{Y})^2 \text{ 最小}$$

(2) 最小二乘公式

根据最小二乘准则推算出的斜率 b 的计算公式为：

$$b = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sum (X - \bar{X})^2}$$

其中 X、Y 分别为样本观测的自变量、因变量取值，$\bar{X}$ 为自变量观测值的平均数，$\bar{Y}$ 为因变量观测值的平均数。

一旦求出了斜率 b，截距 a 就可以根据下面简单的公式求得：

$$a = \bar{Y} - b\bar{X}$$

在女青年受教育年限和收入之间关系的例子中，可以计算出：

$$b = \frac{24473.2}{208.24} = 117.5 \quad a = 1067 - 117.5 \times 9.4 = -37.5$$

因此用最小二乘法拟和出的关于青年受教育年限和收入之



间关系的直线方程为：

$\hat{Y} = a + bX = -37.5 + 117.5X$

如果我们用这个直线方程来预测一个受了10年教育的女青年的收入，只要将 $X = 10$ 代入方程，即可得到 Y 的预测值 $\hat{Y} = 1137$ 元。也就是说，受教育10年的女青年收入大概是1137元。

(3) 斜率 b 的意义

根据定义，直线的斜率等于沿 X 方向向右变化一个单位时，纵坐标 Y 的变化量，即：

斜率 b = X 有一个单位的变化时，伴随着发生的 Y 的变化

$$b = \frac{\Delta Y}{\Delta X}$$

其中 ΔX 和 ΔY 分别表示 X 和 Y 的对应变化量。

在上面的例子中，$b = 117.5$，这就是说，如果女青年的受教育年限增加一年，那么预测她的收入将增加117.5元。

所以 b 值的大小，就是表示 X 对 Y 的影响有多少。b 值有正负之分，如果为正，表示 X 增大，Y 也增大；如果为负，则表示 X 增大，Y 会减少。回归直线的斜率 b 称为回归系数。

3. 决定系数 r^2

如果不借助回归方程，很大程度上我们会用 Y 的平均数来预测未知的 Y 值，此时预测的偏差是 $(Y - \overline{Y})$，而现在利用回归方程和 X 的值来预测 Y 值，预测偏差为 $(Y - \hat{Y})$，利用回归方程预测，偏差会大大减少，因为很大部分偏差 $(\hat{Y} - \overline{Y})$ 由回归解释了，而不能解释的偏差 $(Y - \hat{Y})$ 相对就小得多。那么，总偏差可以分为两部分：

$(Y - \overline{Y}) = (\hat{Y} - \overline{Y}) + (Y - \hat{Y})$

即：总偏差 = 可以用回归解释的偏差 + 不能解释的偏差

可以证明，将这些偏差平方并求和后，类似的等式依然成立，可以推算：

$$\frac{Y\text{的可以由回归解释的变差}}{Y\text{的总变差}}=\frac{\sum(\hat{Y}-\overline{Y})^2}{\sum(Y-\overline{Y})^2}=r^2$$

其中 r 为两个变量 X 和 Y 之间的相关系数。因此相关系数 r 的平方，表示了用线性回归方程作为预测工具时所能减少的变差的比例，如果 r^2 越大，说明回归方程的预测能力越强。因此 r^2 称为决定系数。

在进行回归分析时，最好先计算 r^2 的值，如果这个值很小，那么就不要用线性回归来做预测，因为这样预测误差会很大。

另外值的注意的是，相关系数 r 和线性回归分析都是假定变量 X 和 Y 之间关系具有线性的性质，如果这个假设不满足，就不能用这两种分析方法。

4. 总体回归系数的假设检验

利用样本数据拟和了回归直线，求出了样本回归系数 b 之后，还需要利用它来对总体的情况做统计推断。设总体中自变量和因变量之间的回归系数用 β 表示，可以证明，回归系数 b 近似服从正态分布，其平均数为 β，标准误差 SE 为 $\sigma^2/\sum(X-\overline{X})^2$，其中 σ^2 是观测值 Y 关于总体回归直线的方差，一般是未知的，而用观测值 Y 离开预测值 $\hat{Y}$ 的方差 S^2 来估计。

$$S^2=\frac{\sum(Y-\hat{Y})^2}{n-2}$$

对总体回归系数 β 进行假设检验的零假设和备选假设分别为：

H_0：总体中两个变量没有线性关系（$\beta=0$）



H_1：总体中因变量 Y 随着自变量 X 的变化而线性变化（$\beta \neq 0$）

上述假设是采用双侧检验，如果我们有理论足以支持假设的方向，可以采用单侧检验如下：

H_0：总体中两个变量没有线性关系（$\beta = 0$）

H_1：总体中因变量随着自变量的增加（减少）而线性增加（$\beta > 0$ 或 $\beta < 0$）

对于总体回归系数的检验可以采用 t 检验方法，t 比值定义为：

$$t = \frac{\text{样本估计值} - \text{零假设值}}{\text{样本估计值的标准误差}} = \frac{b}{S/\sqrt{\sum(X - \overline{X})^2}}$$

其中 b 为样本回归系数，自由度为 $n-2$。

5．利用 SPSS 进行回归分析

我们依然使用前面相关分析中，有关研究居民非商品支出与居民文化生活支出的关系的例子，介绍使用 SPSS 进行回归分析的基本步骤。对总体回归系数的零假设和备选假设分别是：

H_0：总体中两个变量没有线性关系（$\beta = 0$）

H_1：总体中因变量 Y 随着自变量 X 的变化而线性变化（$\beta \neq 0$）

- 打开 Analyze 菜单，单击 Regression 子菜单下的 Linear 菜单项，打开 Linear Regression 对话框；
- 对话框左侧为变量名列表框，其中列出了当前数据文件中所有变量的变量名称；
- 对话框的右侧 Dependent 框，用于选定需要进行回归分

析的因变量。将代表月文化生活支出的变量放入其中；

• 对话框的右侧 Independent 框，用于选定需要进行回归分析的自变量。将代表居民家庭非商品支出的变量放入其中；

• Method 框用于选择进行回归分析的方法；

• 单击 Statistics 按钮，可以选择输出需要的统计量；

• 单击 Plot 按钮，可定义输出的统计图，这些图可以帮助分析数据分布的正态性、线性和方差齐次性，还可以帮助检查异常值。

采用 SPSS 系统默认设置输出的回归分析结果见表 11-21：

表 11-21　SPSS 系统默认设置输出回归分析表

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	居民家庭非商品支出[a]		Enter

a. All requested variables entered.

b. Dependent Variable：居民家庭文化生活支出

(a)

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.955[a]	.912	.906	1.5380

a. Predictors：(Constant)，居民家庭非商品支出

b. Dependent Variable：居民家庭文化生活支出

(b)



ANOVA[b]

Mobel		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	392.153	1	392.153	165.783	.000[a]
	Residual	37.847	16	2.365		
	Total	430.000	17			

a. Predictors：(Constant)，居民家庭非商品支出

b. Dependent Variable：居民家庭文化生活支出

(c)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1.105	6.284		-.176	.863
	居民家庭非商品支出	.804	.062	.955	12.876	.000

a. Dependent Variable：居民家庭文化生活支出

(d)

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	72.8761	89.7631	79.6667	4.8029	18
Residual	-2.3300	3.1032	-5.5E-15	1.4921	18
Std. Predicted Value	-1.414	2.102	.000	1.000	18
Std. Residual	-1.515	2.018	.000	.970	18

a. Dependent Variable：居民家庭文化生活支出

(e)

表 11-21（a）是回归分析模型引入或剔除的变量，由于此

例中只有一个自变量，因此引入的自变量是“居民家庭的非商品支出”，方法是强迫进入法（Enter），因变量是“月文化生活支出”。

表 11-21（b）是回归模型摘要，两个变量间的相关系数 $R=0.955$；决定系数 $R^2=0.912$，即引入自变量“居民家庭非商品支出”后，解释了因变量变差的 91.2%，因此可以看出，这个回归模型的解释和预测能力是很高的；调整后的 $R^2=0.906$，由于 R^2 是一个受自变量个数和样本规模之比的系数，有的时候 R^2 倾向于高估自变量对因变量的解释能力，因此需要用调整后的决定系数，才能正确估计；估计值的标准误差 = 1.5380。

表 11-21（c）是方差分析表，显示出回归解释的方差为 392.153，而回归没有解释的方差是 2.365，F 比值 = 392.153/2.365 = 165.783，F 的双侧概值 $p=0.000<0.01$，说明 R^2 统计上非常显著，因此可以认为两个变量之间存在线性关系。

表 11-21（d）是回归分析系数表，显示直线回归方程常数项（截距）= −1.105；回归系数 $B=0.804$，其标准误差 = 0.062，所以此回归直线方程为：$Y=-1.105+0.804X$，即居民家庭的非商品支出每增加 1 元，文化生活支出增加 0.804 元；Beta = 0.955 是标准化回归系数，它是先将自变量和因变量的数据标准化，然后求出的回归系数，此时常数项为 0；回归系数 t 检验的 $t=12.876$，双侧概值 $p=0.000<0.01$，所以回归系数的检验显示在显著度为 0.01 的水平下是统计显著的，即可以拒绝零假设，而认为总体中回归系数 $\beta\neq0$。所以可以认为居民的文化生活支出随着居民家庭非商品支出的变化而变化。如果我们的备选假设为“居民的文化生活支出随着居民家庭非商品支出的增加而增加”，则需要进行单侧 t 检验，单侧



概值可以由双侧概值除以2求得。

表11-21（e）是残差统计表，主要显示预测值、残差、标准化预测值、标准化残差等统计量的最小值、最大值、平均值和标准差等。

第五节　常用统计图表

一、常用统计表

1. 常用统计表

在统计分析中，统计表已经成为基本的表达及显示数据的工具。许多杂乱的统计数据，一经整理在统计表中，就变得清晰易懂，其不同变量或指标间的关系也就一目了然。

基本的统计表我们在前面都已经提到过，如频数表［表11-1（b）］、交互表（表11-16）、相关表（表11-20）等。这些统计表格都是SPSS输出的统计表，而记者在精确新闻报道中制作统计表可以根据需要对于表中的内容和形式进行取舍。

例如，在制作频数表时，文字报道中，说明了样本总数的情况下，直接使用各个项目的百分比，而不再出现频数、累积百分比等项目。表11-22是台湾《联合报》1999年12月21日的精确新闻报道《交通治安整顿成效 不满意比例高》所附的简单频数表。表中给出了1999年年初和年末进行的市民对市府工作成效的满意度调查结果，直接采用满意、不满意的百分比数据对两次调查结果进行了比较，体现了报道的主题——市民对于交通治安整顿成效的不满意比例高，而频数表中没有显示“无意见”或“未回答”人的比例。在表格后面对表中数据进行了说明。

表 11-22 市民对市府工作成效的满意度

市政工作	调查年月	满意	不满意
扫荡色情成效	88/01	55	26
	88/12	44	37
整顿治安成效	88/01	56	22
	88/12	56	36
整顿交通成效	88/01	-	-
	88/12	52	42

注：1）表中数字为百分比，不含无意见和未回答的比例

2）资料来源：联合报系民意调查中心

同样交互表和相关表也是一样，在表 11-16 显示的交互表中，单元格中包含频数、行百分比，列百分比。而记者在进行精确新闻报道时，可以只在制作的交互表中显示能够说明问题的百分比。如表 11-23 所示：[①]

表 11-23 收入与应提倡品质交互百分表

	收入		
应提倡的品质	1000 元以下(%)	1000～2999 元(%)	3000 元及以上(%)
温良恭俭让的品质	12.2	8.1	5.6
四平八稳、按部就班的品质	5.1	4.4	2.8
奋发向上、创新进取的品质	53.2	52.9	58.2
热情乐观、自尊自信的品质	28.3	34.2	30.6
其他	1.2	0.4	2.8
合计	100	100	100

① 范伟达《世纪图景——21 世纪国人生活权威调查》，中国社会出版社，2000 年第 207 页。



2. 制作统计表的注意事项

统计表的形式多种多样，使用时可按照需要灵活设计和选择，但一般统计表都包括4部分，即表头、横行标题、纵栏标题和数据资料。此外，在统计表下得注释、附记、资料来源、某些指标的计算方法等为表外附加的内容。在制作统计表时要注意以下几点：

(1) 每张表都要有号码和标题，标题要简明、扼要，标题一般包含时间、地点、内容，有时可酌情省略。

(2) 注明各种数据的单位，只有一种单位的表格，可在标题中统一加以注明。

(3) 给出必要的说明和标注。

(4) 说明表中数据的来源。

统计表的设计要根据制表的目的，不同的制表目的会有不同的表格与数字的安排。总之，要使统计表简明、紧凑、清晰，使读者能一目了然地看出你所要表达的思想和问题。

二、常用统计图

统计图是统计数据的更直观表现形式，可以将较复杂的数据通过生动的图形表现出来。精确新闻报道中的数据，尤其可以更进一步用图表呈现出来，不仅可以起到美化版面的作用，而且可以吸引读者阅读。以下介绍几种常用的统计图：

1. 条形图

条形图是我们经常见到和用到的一种图形，它用宽度相等的条形的长短来表示各类数据的大小，给人简洁明快的感觉。条形图可以是水平的，也可以是垂直的；其长度可以是频数，也可以是百分比；对于定类和定序变量比较适用。条形图有单

式、复式和分段条形图等多种形式。[①]

（1）单式条形图

单式条形图一般用来表现一个变量的频数或百分比。如图11-6所示。从图中很明显可以看出，多数人认为“网络”将是21世纪人们信息的主要来源。

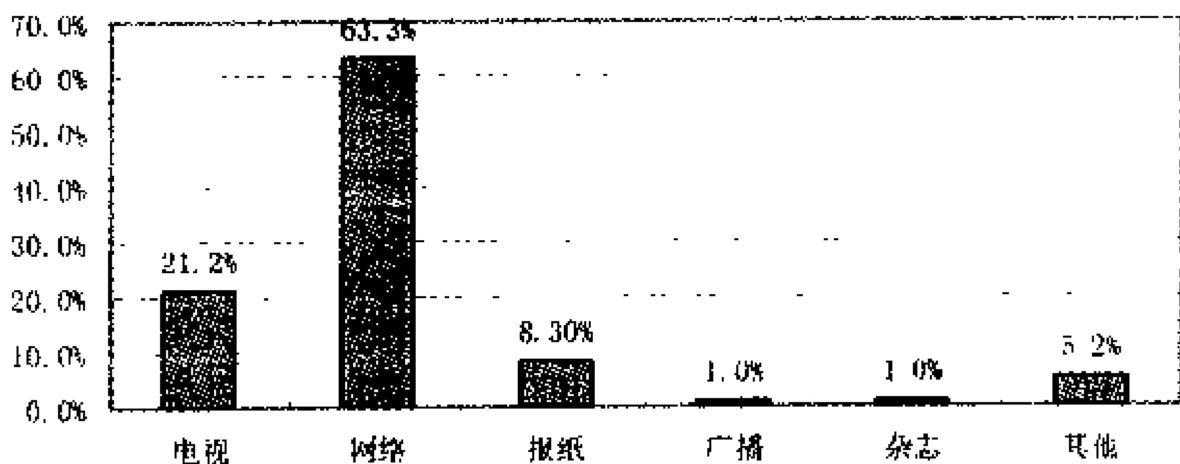


图11-6　21世纪人们信息的主要来源

（2）复式条形图

复式条形图一般用来表现两个变量交互分析的数据结果，以两个或两个以上条形为一组进行比较，既可以比较组内的条形高度，又可以比较组间的差异。如图11-7。

图中显示学历越高，认为“经济困难”是对未来几代人最大威胁的比例越低；1999年调查中，中、低学历人认为“经济困难”是对未来几代人最大威胁的比例比1997年高出10个百分点。

① 袁卫、吴喜之、贾俊平《描述性统计学》，中国统计出版社，1996年，第79页。

认为经济困难是对未来几代人的最大威胁①

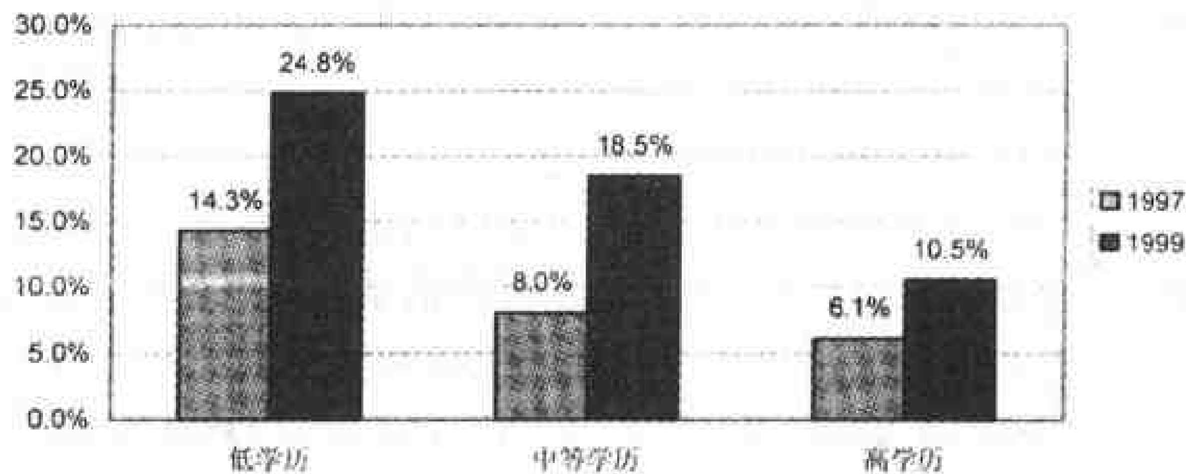


图 11-7　不同学历的人认为“经济困难”是对未来几代人最大威胁的比例

(3) 分段条形图

分段条形图通常是以条形内部的分段表明总体内部构成的情况。如图 11-8 所示：

是否喜欢现在居住的城市②

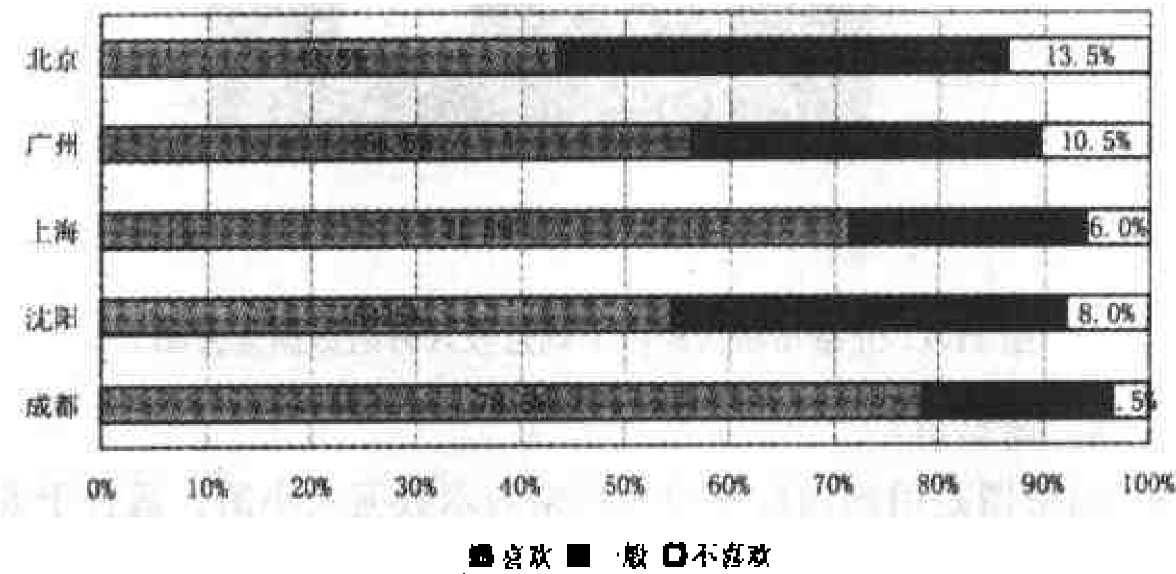


图 11-8　是否喜欢现在居住的城市

①② 资料来源：1999 年《IEM30 国世界环境意识调查》中国调查报告。

图中可见，成都人喜欢自己现在居住城市的比例最高，其次为上海人，而北京人喜欢自己现在居住城市的比例是几个城市中最低。

2．直方图

直方图表征数据的频数分布特征，与条形图类似，都是用条形长度来表示数据特征，但直方图中的条形之间是没有间隔的。而且直方图适用于表现各种测量级别的变量。图 11-9 是使用 SPSS 绘制的北京市居民家庭平均月收入分组数据直方图。

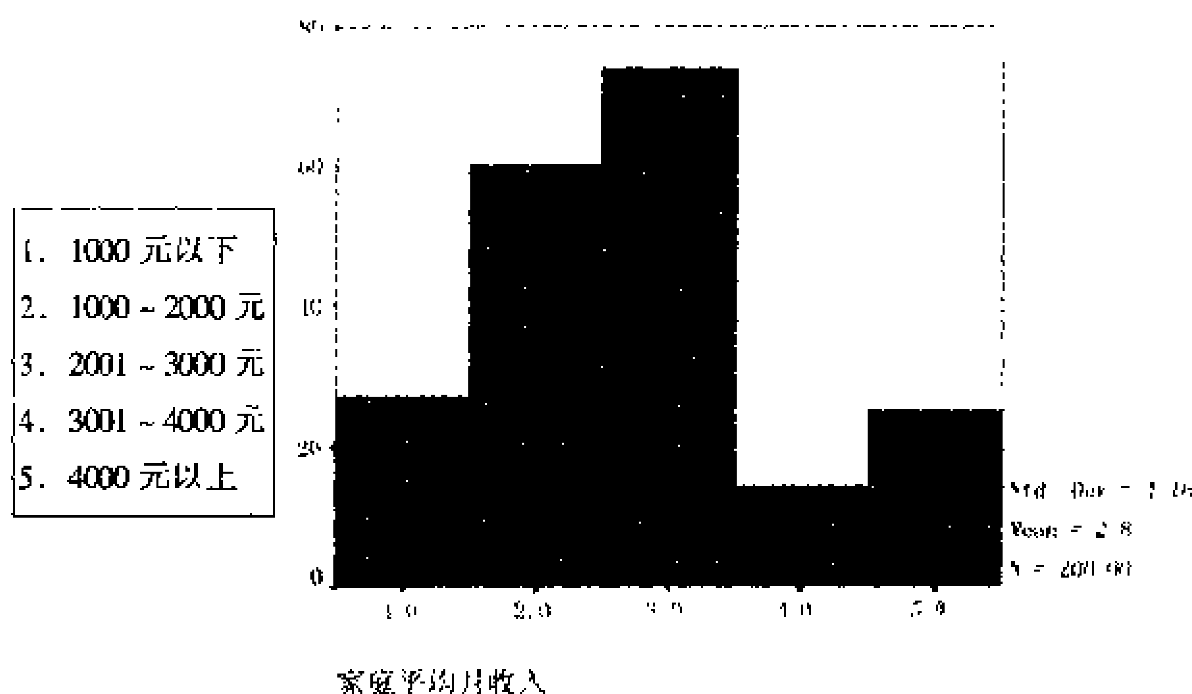


图 11-9　北京市居民家庭平均月收入分组数据直方图

3．圆形图

圆形图是用圆内扇形的面积来表示数据大小的，适合于表现单选问题，整个圆形面积总计 100%，每一部分的面积就表示了某个变量对应取值的百分数。圆形图可以是平面的，也可以是立体的。如图 11-10 所示。图中显示了被调查者的学历结构分布，可以看到中等学历的被调查者几乎占了半数。



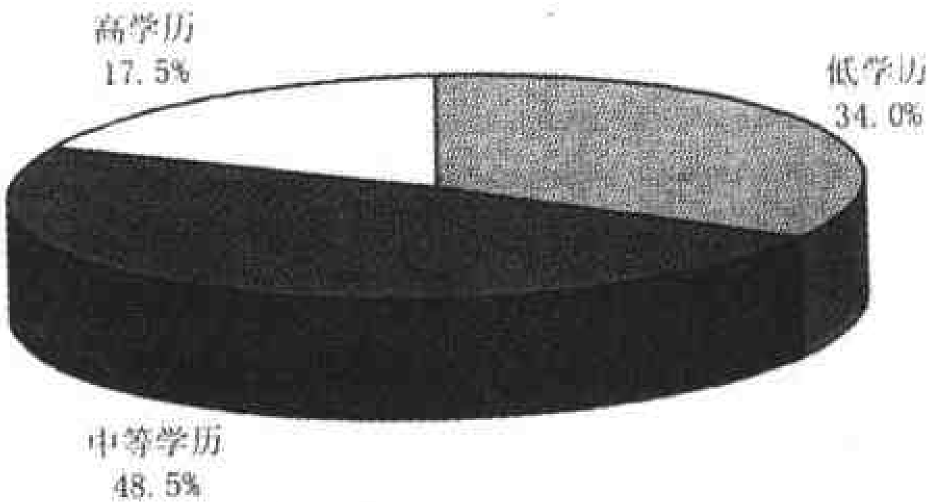


图 11-10 被调查者的学历结构

4. 线图

线图是在平面坐标上用折线表示数量变化规律的统计图。主要描述变量的取值在不同项目上的变化，如表现时间序列的数据—历年的出生率、死亡率等等，或者表现被调查者在各个项目上的态度得分，或对不同对象的评价得分等等。图 11-11 是折线图的一个例子。这是某地居民对于当地交通状况历年的评价情况，评价采用五级量表，其中 5 表示“非常满意”，1 表示“非常不满意”。

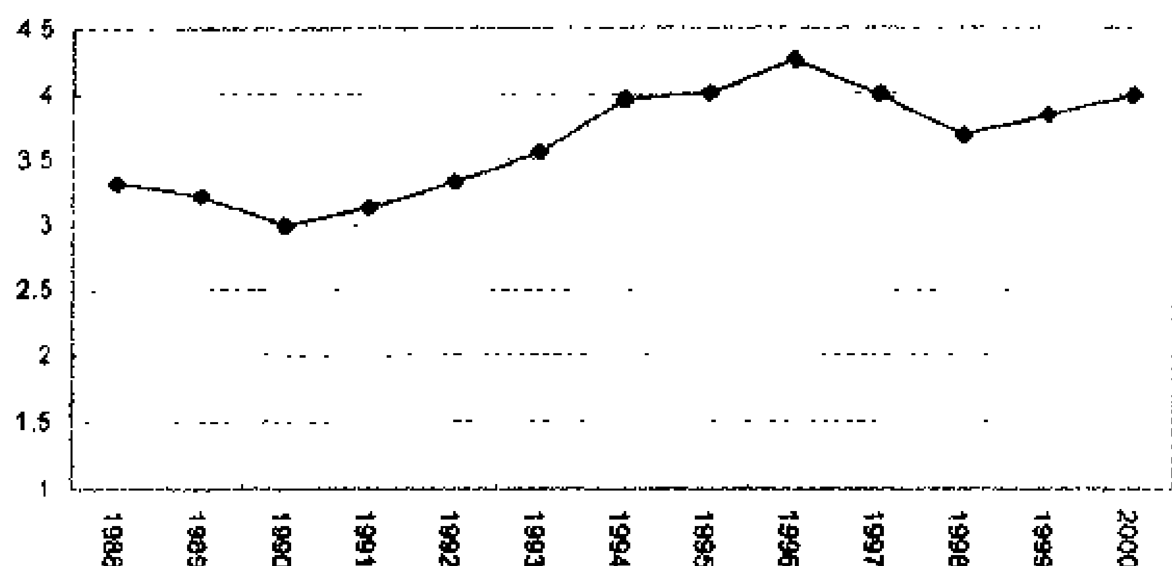


图 11-11　某地居民对当地交通状况的评价

5. 散点图

二维散点图是利用两个变量的取值为坐标，在坐标系中描绘出确定的点，然后通过这些点的分布特征来显示数据分布特征。利用这种散点图，可以了解两个变量之间的关系。例如，我们在前面第四节中利用散点图（图 11-4）考察两个变量间的相关分析。图 11-12 是用前面提到的女青年受教育年限和收入的数据制作的散点图。

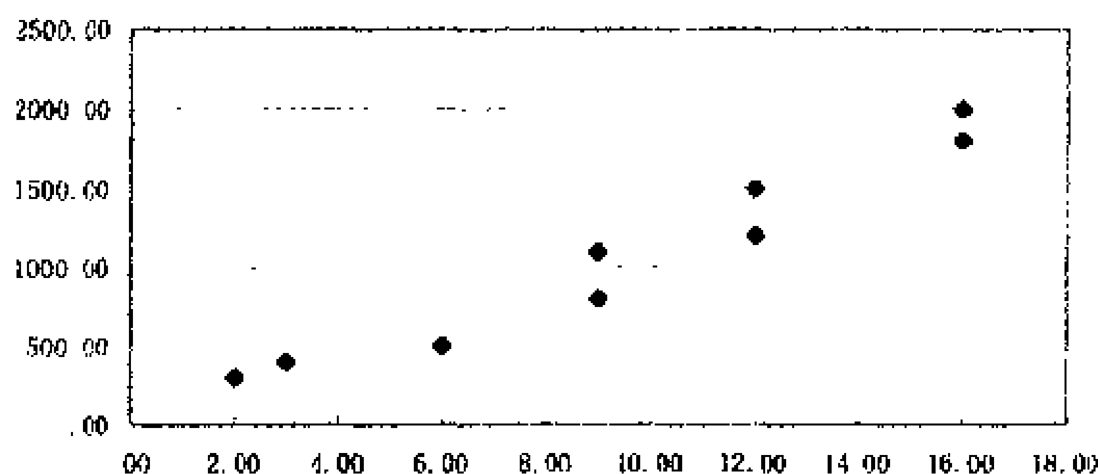


图 11-12　女青年受教育年限和收入关系的散点图



前面介绍的多是最基本的统计图，现在很多计算机软件都有很强的绘图功能，可以将图形绘制的十分美观、生动。但是记者在试图绘制统计图的时候，一定要注意将图绘制得标准、规范，从指标的选择、图形的创意、绘图的准确等方面都要认真推敲，否则，可能会给读者造成错觉。

第十二章　精确新闻报道的写作技巧

在对收集到的数据进行分析之后，就要向受众报告研究结果，在这里，我们将介绍两种基本的精确新闻报道的写作模式，即精确新闻倒金字塔式写作模式和精确新闻钻石式写作模式。

第一节　精确新闻倒金字塔式写作模式

一、倒金字塔式写作模式

倒金字塔式是新闻界最传统的写作模式，这种模式最早出现在美国内战时期，至今已有100多年的历史，但仍然是目前新闻界较为常用的写作模式。

这种模式的最大特色，是把最重要、最新鲜、最精彩的内容放在新闻的第一部分，把次重要的内容放在第二部分，新闻的其他内容则依照内容重要性递减的原则排列。采用这种写作模式的新闻，在形式上就像一座倒立的金字塔，因此被称为倒金字塔式写作模式。

倒金字塔模式能在新闻界盛行百年而历久不衰的原因，是因为这种写作模式能满足新闻工作及受众的需要。



首先，倒金字塔式写作模式能便利报纸编辑作标题、安排版面，便利广播电视编辑制作内容提要和节目排序。

在编辑制作报纸标题或广播电视新闻内容提要时，由于记者把新闻最重要的内容放在前面，编辑只要阅读或视听新闻的前几段，便能掌握新闻重点，便于编辑制作标题或内容提要。如果记者写稿或制作节目时，不采取倒金字塔式写作模式，而把新闻的重点分散在各段、各部分中，编辑可能需要阅读完全文或视听完全部节目，直到找出新闻的重点后，才能制作标题或内容提要。这样一来，编辑工作的时效性将难以掌握。

在安排版面或栏目时，由于报纸的每一版或广播电视的每一档节目能容纳的新闻有限，编辑常需要删减若干新闻的篇幅或长度，才能顺利排版、排序。倒金字塔式写作模式把新闻最不重要的内容放在最后部分，使编辑能在不影响新闻主要内容的情形下，轻易删掉不重要的内容。如果记者采取别的写作方式或节目制作方式，删减工作将难以进行，因为编辑必须把全文改写或重新编辑制作后，才能缩短新闻的长度，并保留新闻的重点。

对受众来说，倒金字塔式开门见山的写作方式，比较能够符合现代受众获取信息的需求。因为报纸的张数越来越多，广播电视新闻节目也越来越多，受众的读报、听广播、看电视的时间有限，通常无法获取所有的新闻。因此，受众获取新闻时，可能会对媒体上的新闻加以选择，也可能在获知一篇报道的前几部分后，而选择其他的新闻。

倒金字塔式写作模式把最重要的内容放在新闻的最前面，能增加新闻对受众的吸引力，也能帮助受众选择新闻。此外，在受众选择新闻时，只要获知一篇报道的前几段，便能使他们对这篇报道有大致的了解，不至于对所发生的事情一无所知。

除此之外，倒金字塔式写作模式也符合人类传递重要或紧急事件时的传播习惯。以北京市民大力支持北京申奥活动这个事件为例，人们在传递这条消息时，很少有人会从北京申奥的目的、北京申奥的优势和困难、北京奥申委的宣传活动等开始谈起，再慢慢谈到北京申奥活动得到北京市民的大力支持。绝大多数的人在谈话时，即开门见山地指出最重要的事实：北京市民对申办2008年奥运会的支持率达到94.9%。

正因为倒金字塔式写作形式具有以上长处，所以从19世纪末一直沿用至今。

当然，倒金字塔式写作模式也不是没有缺点，最大的缺点，是这种模式把新闻的重点放在第一部分中，编辑作标题或提要时，又可能把相同的内容放进标题或提要中，使标题、提要与新闻第一部分的内容重复，受众获取信息时，必须重复获知相同的内容。

此外，这种模式强调固定的形式，无形中限制了记者的创造力，记者的新闻表现技巧也不易发挥，因而使新闻的可读性、可听性、可视性受到影响。

二、倒金字塔式在精确新闻写作上的应用

1. 倒金字塔式导语写作

在倒金字塔式写作模式中，导语是新闻的第一部分，也是最重要的一部分。好的导语不仅要呈现新闻事件最重要的内容，也要能引发受众的进一步了解新闻事实的兴趣。通常来讲，一则好的导语应具备4项重要功能：

（1）简明扼要地指出新闻的内容，清楚呈现新闻的要点；

（2）迅速使受众对新闻的内容感到兴趣；

（3）为编辑选择稿子、尽快决定稿件的取舍及制作标题等



提供了方便；

(4) 使受众获知新闻更具效率，因为每一篇报道的重点，都在开始时即已指明，受众不必浪费时间获知不感兴趣的新闻。

由于导语在新闻中地位特别重要，倒金字塔式写作模式强调导语应尽可能地包括新闻中的6个要素：何人？(Who)；何事？(What)；何时？(When)；何处？(Where)；何故？(Why)；如何？(How)。

在倒金字塔式的写作模式中，这6个要素（5个W和1个H）被认为是新闻最重要的内容，应该尽可能地放在导语中。但在实际中，并不是每一篇报道都需要把这6个要素全部放进导语中，因为如果把这6个要素全部放入导语中，可能使导语变得过分冗长，而影响受众的兴趣。为了使导语能容纳新闻中最重要的内容，又不至于过分冗长，新闻界通常采取的办法，是在写作或制作新闻前，先把这6个要素依重要性加以排列，在写作过程中，再从这6个要素中选择最重要的几个要素放在第一部分中，其他的要素则依重要性递减的原则，分散放在新闻的第二部分或第三部分中。

在现代的倒金字塔式写作模式中，导语不仅是指新闻的第一部分。有的时候，新闻的第二部分，甚至第三部分都可能被认为是一篇报道的导语。

2. 精确新闻报道的导语写作

(1) 导语的重点

在传统的倒金字塔式写作模式中，5个W及1个H被认为是新闻中最重要的内容。但就精确新闻报道而言，新闻最重要的内容，应该是研究的结果，而非5个W和1个H。

举例而言，盖洛普（中国）咨询有限公司在北京提出申办

2008年奥运会后，进行了一项民意调查。这项民意调查，主要探讨下列3个问题：

①北京市民对北京申办奥运会的支持率有多少？

②北京市民对北京申奥成功是否报有信心？

③北京市民对北京申奥活动了解程度如何？

对于这样的一项精确新闻报道研究，如果写作时仍然强调5个W和1个H并不是新闻的最重要内容。在盖洛普公司这项研究中，5个W及1个H所代表的意义如下：

①何人：北京市民

②何事：进行北京市民对北京申奥活动的民意调查

③何时：民意调查进行的时间

④何处：北京市

⑤何故：这是北京第二次申办奥运会，希望了解社会北京市民对申奥的态度

⑥如何：如何进行调查

因此，强调5个W及1个H的写作格式并不适用精确新闻写作。记者在采写精确新闻时，必须放弃强调5个W和1个H的传统写作格式，而采取强调研究结果的写作格式。

写作精确新闻的通常做法是：先把所有的研究发现列出，然后根据记者编辑个人的判断，把各项发现依重要性加以排列。以上述盖洛普公司的调查为例，就可依照重要性将3项发现加以排列。

最重要：北京市民对北京申奥的支持率如何。

第二重要：北京市民对北京申奥成功是否报有信心。

第三重要：北京市民对北京申奥活动的了解程度如何。

把研究发现依重要性排列后，便可依照这个顺序选择导语应包括的内容。



（2）导语的形式

精确报道新闻的导语和一般倒金字塔式新闻导语一样，可以采用许多不同的形式。下面是5种最常见的导语形式：

①摘要式导语

摘要式导语是把最重要的调查研究发现，以摘要的方式放在导语中，简明扼要地写出来。这种形式的导语是精确新闻报道最常见的导语形式。

本报民意调查发现，在高雄市各项市政议题中，民众满意度最高的是环保工作，满意比率达六成九；对于自来水质的满意度虽增加了五个百分点，但仍有八成市民不满意。

（《联合报》，1988.12.21）

上面这则导语以摘要的方式，列出《联合报》这项民意调查的两项最重要发现，调查中的其他重要发现，则按重要性，依序分散放在这条新闻的后面几段中，以避免使导语过分冗长，而影响导语的可读性。

这个例子也说明了，写摘要式导语时，如果所进行的精确新闻报道研究，只有一项或两项重要发现，记者当然应该把这一项或两项重要发现纳入导语中。但如果所进行的精确新闻报道研究有多项重要发现时，记者就必须从所有的重要发现中，选择一两项最重要的发现放在导语中，并把其他的重要发现依重要性递减的原则，分散放在新闻的各段中。这样才能使导语精简扼要，引起受众兴趣。

②对比式导语

当精确新闻报道研究发现，民众对所探讨的人物与事件同时具有正面与反面两种对立的态度，或是存有矛盾态度时，便

可以采用对比式的导语，来突出这种对立的态度与矛盾的情结。

12月11日下午，7位外国飞行勇士和1对新婚夫妇乘坐飞机终于安全通过了张家界的瞩目景观天门洞。人们对在世界自然文化遗产进行这样一种冒险，报有一种矛盾的心理。本报的民意调查显示，55%的被访者认为如果能增加张家界国际上的知名度，带来相应的旅游收入，冒一点风险是值得的；余下45%的被访者则反对拿世界认可的宝贵的自然文化遗产冒险，认为这种活动就不应该举行。

（《北京青年报》，1999.12.15）

③背景式导语

有的时候，记者会觉得有必要在新闻的开头，先说明进行精确新闻报道研究的背景原因，才能使受众更清楚地了解研究发现的意义。

苏俄乌克兰车诺比尔（即切尔诺贝利，编者注）核能电厂于4月底突然爆炸，据外电报道已造成一万余人死亡的惨剧，辐射尘已扩散到欧、亚、美等洲，引起全世界多数民众的恐慌。

根据本报对台北地区民众的一项调查显示，苏俄核电厂事件后，在台湾也引起民众的“车诺比尔事件恐惧症”，38%的人认为核能发电有危险，42%的人不知道核能电厂究竟安不安全。

（《联合报》，1986.5.12）

④第二人称式的导语

这种形式的导语是在导语中采用第二人称——你，来描述研究发现，使受众与研究发现间产生某种关联性。

如果你是本区典型的居民，今年你将看5次医生。如果你的医生也是本区的典型医生，他将没有时间为你提供你希望提供的医疗服务。

由9个郡所组成的都布魁医疗服务区，共有25.2万位居民，但只有177位忙得不可开交的医生。

（《都布魁电讯前锋报》，1975.8.17～24）[①]

⑤问句式导语

另外的一种常见的精确新闻报道的导语写作形式，是问句式导语。这种形式的导语，如果运用得当，可以激发受众的好奇心，引起受众的阅读欲望。

近日，酝酿已久的利息税终于出台了。当利息税这块石头一下子投入居民储蓄这潭静水之后情况如何？据本报的调查表明，32%的被访者认为，国家征收利息税对他们的家庭收入“没有影响”，57%的被访市民虽然肯定这种影响的存在，但也表示对其家庭收入的影响是“较小”的。

（《北京青年报》，1999.9.8）

(3) 新闻的结构

精确新闻报道采用倒金字塔式写作模式时，除了应在导语内写出最重要的内容外，新闻的其他段落必须要依照重要性递

① 资料来源：罗文辉《精确新闻报道》，正中书局，1997年，第294页。

减的原则加以排列。

一般来说，第二段应该呈现无法纳入导语的重要内容，或是对导语的内容提出更进一步的说明，其他各段则应逐段描述各项发现。在安排各段的内容时，每一段最好只描述一项发现，以遵守新闻一段一主题的原则。

此外，段落与段落之间也应该使用适当的转接句，使段落间的转接顺畅。

北京的申奥活动得到北京市民的大力支持

据盖洛普（中国）咨询有限公司的一项调查表明，北京市民对申办2008年奥运会的支持率达到94.9%，这个数字大大高于巴黎79%和多伦多78%的支持率。

调查表明，北京市民对申奥的这种支持带有强烈的普遍性，不同群体（性别、年龄、户口所在地）对申奥的支持率不存在差别。其中，北京男性居民对申奥的支持率为94%，女性支持率为96%，城镇居民和乡村居民的支持率同为95%，18岁至30岁人群、31岁至60岁人群、61岁以上人群的支持率分别为95%、94%和99%。可以说，在北京，无论城镇还是农村，无论男女老幼，都表示了全力支持北京申奥的强烈愿望。

在另外5.1%没有表示支持北京申奥的被访者中，有4.4%的人表示了无所谓的态度，0.4%的人拒绝回答，只有0.3%的人表示反对。

负责此项调查工作的吴涛博士说："在我接触过的调查中，从来没有见过像94.9%支持率这样高的数据，一般国外对于政府某项活动的支持率，大多也只是在50%到60%左右。"

另外，在支持北京申奥的北京市民中，有93%的人对北京申奥的成功抱有信心，这也是一个非常高的信心指数。值得



注意的是，在成功信心这组数据中，尽管有93%的人对北京申奥的成功抱有信心，但表示“比较有信心”的市民只占到31%左右。这说明，北京人比1993年第一次申办奥运时更加成熟、冷静，回答问题有所思考和保留。

北京居民对北京申奥相关内容的了解程度如何呢？调查发现，尽管在知道北京申奥的北京市民中，有近93%的北京市民知道北京申办的是2008年奥运会，但能正确答出北京申办的是“第29届奥运会”的人只有24%；另外，只有43%的人知道申奥的会徽，只有40%的人能说出北京申奥口号“新北京，新奥运”。

为了解北京市民对北京申办2008年奥运会的态度，盖洛普（中国）咨询有限公司受北京奥申委的委托，从2000年11月8日至11月22日对1626位18岁和18岁以上北京市民进行了电话访问或入户调查。调查采取了随机抽样的方式，抽样误差在3%左右。[①]

上面这条新闻是精确新闻报道采用倒金字塔式模式的例子，这条新闻共有7段，第一段的内容包含了这项民意调查最重要的一项发现。因此，这一段是这条新闻中最重要的一段。

这条新闻的第二段、第三段和第四段，是对第一段中的重要发现进行更进一步说明，使受众对这项调查及研究发现有更深一层的了解。可以说，第一段到第四段是整条新闻的核心。

这条新闻的第五段描述第二重要的研究发现，第六段则描述第三重要的研究发现。这二段在整条新闻中的重要性比不上前四段，因此被安排在新闻的中后部分。而这二段的先后顺

① 数据来源：《北京青年报》网络版2001年2月20日。

序，也按照重要性递减的顺序加以排列。

这条新闻的最后一段，说明这项民意调查的过程与方法。把研究过程与研究方法方面的信息，安排在新闻的最后一段，并不表示这些内容是新闻中不重要的内容。相反，这些信息是每一则精确新闻报道都应具备的内容。

第二节 精确新闻钻石式写作模式

一、钻石式写作模式

1. 典型的钻石式写作模式

虽然国内新闻写作的教科书很少介绍钻石式的写作模式，但新闻界采用这种写作模式已经有很长的历史。尤其是在国外的新闻报纸或新闻杂志上，经常可以看到采用这种写作模式撰写的深度报道或特写。

钻石式写作模式与倒金字塔式完全不同，这种写作模式并不需要把最重要的内容放在导语内，也不需要把新闻的段落按照重要性依次排列。

钻石式写作模式至少应包括导语、主题、主体及结尾4个部分。

钻石式写作模式的导语，呈现的是人物或事件的现状，所描述的可能是一段感人的情节，也可能是个人的独白等等，其目的是提出生动的实例，以引发受众了解新闻事实的兴趣。

主题的作用在揭示全文主旨，并在导语与主体之间扮演承先启后的桥梁角色，使导语中所描述的情境，能流畅地与主体的内容衔接。

主体所呈现的是人物或事件的历史背景，展现的是从遥远



的过去，逐步发展到现在的情况，这样才能使全文在结尾时转回导语中所描述的情境。

结尾是钻石式写作模式的重点，钻石式写作模式强调写作应首尾连贯。如果一篇报道以某一位人物开始，结尾也应以这一人物收尾；如果以一段戏剧性的情节开始，结尾时也要返回这段情节才能收场，这样，才能使前后呼应，始终如一。

由于钻石式写作模式特别重视首尾连贯，结尾不仅需要和导语紧密结合，也应该是全文最精彩的部分，因此，当编辑遇到版面不足，栏目时间容纳不下，需要删减报道的长度时，对采用钻石写作模式采写的报道，不能从报道的后面开始删减，必须设法在报道的主体部分加以删减，因为主体部分通常呈现的是人物或事件的历史背景，稍加删减，将不至于对新闻报道的完整性产生太大的影响。

2．华尔街日报公式

在全球各地著名的报纸中，把钻石式写作模式应用得最为成功的是《华尔街日报》。该报定期采用钻石式写作模式来探讨美国及全球性的问题，并规定刊登在该报的第一栏上。

由于《华尔街日报》把钻石式写作模式运用得极为成功，并且逐渐建立了一种独特的风格，因此一些美国新闻界人士把《华尔街日报》独特的钻石式写作模式称为华尔街日报公式。

华尔街日报公式可以看做是钻石式写作模式的一种变体，这种写作模式的基本观点是：一千万人死亡只是一项统计数字，一个人死亡却是一场悲剧。

对多数受众而言，失业率、国防预算、财经问题及国际冲突等，只是一些复杂难懂而又与自己无关的问题。但如果记者能从个人的角度出发，描述或解释这些问题对个人可能产生的影响，不仅能使这些复杂的问题变得简单易懂，也很可能使受

众产生共鸣，激发受众进一步了解的欲望。

因此，华尔街日报公式认为，失业率高达7%，对多数受众而言并无意义，但如果我们介绍一家工厂因经济不景气而被迫关门，导致一位失业工人全家生活陷入困境，然后再谈论失业问题，就比较能使受众了解失业率的意义，以及失业率的数字可能引起的对社会的冲击。

使用华尔街日报公式进行新闻报道时，共有4个步骤：第一步骤，集中描述个人的情况，如失业工人个人全家的困苦情形，这是华尔街日报公式最重要的一个步骤；第二步骤，设法在个人与大的问题（如失业率、国防预算等）之间建立联系，使报道能从个人的问题，转接到大的问题上；第三步骤，对大问题的报道，如当前社会的失业率状况，国防预算的分配、审核情形等；第四步骤，需要回到第一步骤中所描述的个人情境，并设法以感人或有力的情节收场。

二、钻石式在精确新闻写作上的应用

目前，国内从事新闻报道的记者，很少有人尝试采用钻石式写作模式。记者如果要把钻石式写作模式应用在精确新闻专题报道上，除了需要进行精确新闻研究外，还需要用传统的写作方法加以配合。

美国《芝加哥论坛报》曾在1985年用内容分析方法分析了法院的记录，探讨芝加哥保释制度的不公平现象。该报根据这项内容分析的结果，发表了一系列报道，其中一篇就采用了采访与调查并重的钻石式写作模式。下面是这篇报道的节选：[①]

---

① 资料来源：罗文辉《精确新闻报道》，正中书局，1997年，第302页。



3年前6月的一个星期四晚上，原本是一场生日宴会，却变成了黎佛拉的痛苦经历。他因为付不出保释金，而在库克郡监狱度过了12天。

黎佛拉当时只有19岁，在一家餐厅工作，他因一项他不曾犯过的罪而被捕入狱。黎佛拉的例子显示芝加哥的保释制度，有严重的不公平现象。

论坛报用电脑进行了一项保释制度的研究，结果发现每年有13000人被逮捕入狱，但法律却无法证明他们曾经犯罪。其中有3000人被监禁的时间超过两个星期，有些人入狱的时间甚至超过一年。

这些法律认为无罪的人，却遭受了比成千上万的重刑罪犯更严重的惩罚。成千上万的重刑罪犯，在审判前被保释在外，罪行确定后，又获得缓刑。

论坛报的这项研究，仔细分析了1983年至1984年13个月期间，在芝加哥发生的2759件案子。结果发现，每年无辜的民众因付不出保释金而被关入监狱，其中7327位被告甚至没有刑事前科。

黎佛拉在1983年6月3日被监禁前并无前科，他从3岁起便住在芝加哥，他和他父母一起住在一栋公寓内，并且已经订婚。

他说，在被关的12天里，他两度被看守殴打，并且失掉了他的工作……

以上的例子所采用的写作模式就是钻石式写作模式。这篇报道基本上分为4个部分。

第一部分：这篇报道的前两段描述的是黎佛拉的不幸遭

遇，主要目的是以生动的实例引发受众的阅读兴趣。这两段符合钻石式写作模式写作导语的要求，也与华尔街日报公式强调第一步骤应集中描述个人情境的论点不谋而合。

第二部分：这篇报道中，第二段的最后一句指出："黎佛拉的例子显示芝加哥的保释制度具有严重的不公平现象。"这一句话点出了全文的主旨，发挥了钻石式写作模式中主题的功能。此外，这个句子也把全文讨论的问题由导语中描述的黎佛拉个人遭遇成功地转接到大的问题（芝加哥的保释制度）上，等于完成了华尔街日报公式的第二个步骤。

第三部分：从第三段到第五段，这篇报道开始逐步介绍《芝加哥论坛报》所进行的这项研究及研究发现。这三段是钻石式写作模式的主体，也是华尔街日报公式的第三步骤：对大问题的报道。

第四部分：这篇报道的最后两段从芝加哥的保释制度转回到黎佛拉的个人遭遇，使结尾和导语前后呼应，这种做法满足了钻石式写作模式首尾连贯的要求，也完全符合华尔街日报公式第四步所要求的：回到导语中所描述的个人情境。

上述例子显示，钻石式写作模式是极具震撼力写作方式，这种写作模式把传统采访写作触动人性的潜在的潜能与精确新闻报道的科学精神溶为一体，使精确新闻报道得以发挥更大的冲击力。

但钻石式写作模式要求记者根据研究发现选择适当的采访对象进行访问，记者在采访时，可能不容易找到适当的采访对象，有时候，即使能够找到适当的采访对象，访问工作也可能极为耗费时间，因而迫使精确新闻报道必须延后发布。因此，钻石式写作模式可能不适合用于少数特别重视时效的精确新闻报道研究上。此外，采用钻石式新闻写作模式时，还需要注意



下列几点：

1. 新闻机构必须能够提供较多的篇幅，因为采用钻石式写作模式所写出的文章通常比较长，如无足够的篇幅，不宜采用这种形式的写作模式。

2. 钻石式写作模式极为耗费时间，新闻机构应多给记者足够的时间进行采访及撰稿。

3. 钻石式写作模式比较适合用于特写及深度报道。不适合用于一般的消息写作。

4. 杂志记者比报社记者更适合采用钻石式的写作模式，因为杂志出刊时间间隔较长，也有足够的篇幅刊登较长的文章。

第三节　精确新闻基本形态及写作中应注意的问题

一、精确新闻报道的3种基本形态

目前，精确新闻的报道形态基本有3种：消息、专稿或专栏、消息配合专稿或专栏。

1. 消息。对主要调查结果进行简短、迅速、准确报道，对调查结果不作详细分析、说明，强调新闻的时效性。

2. 专稿或专栏。不仅呈现调查的结果，而且还以特写、深入报道或新闻分析等其他方式对调查结果加以分析、评述。

3. 消息配合专稿或专栏。以上述消息及专稿或专栏两种或两种以上形式出现。

在报纸、杂志以及互联网上，以上3种精确新闻报道的形态都可配以图表对调查结果进行说明，这些图表包括表格、图

形、漫画等，这种方式有利于受众对调查结果的直观理解，对数字结果产生形象化概念。

前面我们介绍的两种基本的精确新闻写作模式，其中倒金字塔式写作模式更多地用于消息写作，而钻石式写作模式则比较适用于写深度报道。记者在进行精确新闻写作时，并不一定要采用这两种基本模式，可以按照实际需要灵活选择写作形式。但是不管采用哪种写作形式，有些问题是必须要考虑，并加以注意的。

二、精确新闻写作中涉及研究过程与方法的7个必要内容和注意事项

与其他类型新闻相比，精确新闻的优势和特点在于它的数字化和精确性，受众对精确新闻可信度的要求也很高。因此，在精确新闻写作中，有关调查过程与方法的7个方面的内容一定要在新闻中有所体现，以加强精确新闻报道的可信度。这7个方面的内容分别是：

1. 受何机构委托进行调查，调查的目的是什么

作为新闻媒体，在发表民意调查结果时，一定要指出是哪个机构委托进行的调查，目的是什么。因为，委托机构的目的可能会使调查有一定倾向性，可能会设计一些带有引导性的问题或采用代表性比较差的抽样，这会对调查结果的公正性有一定影响。一个明显的带有倾向性的调查结果的发布，会影响到媒体的声誉和自身的形象。新闻媒体要把委托者的目的讲清楚，以体现新闻的客观性，为受众理解调查结果提供必要的参考。

2. 调查的执行机构是谁

在精确新闻中，一定要明确民意调查是由哪家机构完成



的。是专业市场调查公司、学术研究机构、大众媒体本身、政党或其他团体完成的。调查机构本身的声誉对于判断民意调查结果是非常有帮助的。

3．共对多少人进行了调查，有效样本是多少

调查10个人所得出的结论，与调查1000个人所得出的结论的分量是不同的。在其他条件相同的前提下，样本量越大，有效样本越多，调查结果的准确性也就越高。因此，在精确新闻报道中，这两点对调查结果可信度而言是必要的。

4．被调查者是如何抽选出来的

一般来说，民意调查的目的是反映普遍的民意，但由于条件限制，往往不可能对所有人进行调查，只能采取抽样的方式。被抽选出来的部分人是否对全部人群具有代表性，他们的意见是否能够代表普遍意志？这是受众十分关心的问题。所以，在精确新闻报道时，一定要说明抽样方式和被抽样人群的总体特征，以体现被调查者的代表性和调查结果的普遍性。

5．抽样误差是多少

样本是总体的一部分，虽然有代表性，但是并不等于总体。用样本估计总体肯定会有一定误差。抽样误差量越小，估计量的精确度越高。因此，我们在精确新闻写作时，一定要说明这次调查结果的抽样误差是多少，否则，受众会对调查结果的准确性产生怀疑。

6．调查所采用的访问方式是什么

我们知道，调查方式主要有电话访问、入户访问、邮寄访问、拦截访问等多种方式，各种不同的访问方式各有其利弊和一定的局限性。因此，精确新闻写作者有必要让受众了解调查所采取的访问方式。

7. 什么时间实施的调查

在精确新闻中，明确告知调查实施的时间，对受众理解调查结果也是很有益的。另外，调查实施的时间对具有时效性的问题是非常重要。因此，在精确新闻报道中，调查实施的时间应有所体现。

对精确新闻而言，在报道中对以上7个内容的必要解释说明是不可或缺的，这也是精确新闻写作相对于其他类型新闻写作的特点之一。

我们前面提到的“北京申奥活动得到北京市民的大力支持”这条新闻的最后一段，基本体现了精确新闻报道对这7个内容的要求。

为了解北京市民对北京申办2008年奥运会的态度，盖洛普（中国）咨询有限公司受北京奥申委的委托，从2000年11月8日至11月22日对1626位18岁和18岁以上北京市民进行了电话访问或入户调查。调查采取了随机抽样的方式，抽样误差在3%左右。

这段文字表达了以下内容：本次调查是受北京奥申委的委托进行调查，调查的目的是了解北京市民对北京申办2008年奥运会的态度；调查的执行机构是盖洛普（中国）咨询有限公司；共对1626位北京市民进行了调查，这些市民是采取随机抽样的方式选出的，共同特征为18岁和18岁以上北京市民；抽样误差在3%左右；调查实施的时间为2000年11月8日至11月22日。

另外，前面我们也曾经提到，在问卷中相同内容的问题也会因为措辞的不同，而导致不同的调查结果。因此美联社执行



编辑协会（The Associated Press Managing Editors Association）曾建议新闻界在报道调查结果时，应该列出问卷中所有的问题，以供读者参考。但也有些学者认为限于版面和新闻价值等问题，不必列出全部的问卷问题。

在精确新闻写作中，涉及调查过程与研究方法方面的内容还应注意下面两点：

第一，除上述7个有关研究方法与研究过程的重要内容外，其他艰涩的统计及研究专用术语，如研究假设、变量、操作性定义、卡方分析、回归分析等，应尽量避免出现在新闻中。

第二，资料分析求出的统计数字，如百分比或平均值等，在新闻中应尽量采用整数。统计数字不是整数时，应依据四舍五入的原则处理成整数，以方便受众。

参 考 文 献

《北京青年报》、《北京晚报》、《深圳特区报》、《环球时报》、《联合报》、《市场调查》(全国市场研究行业协会出版)、《现场统计与管理》

[美]艾尔·巴比	《社会研究方法》	华夏出版社	2000年5月
[美]伊恩·罗伯逊	《社会学》	商务印书馆	1994年
Herbert Asher	《Polling and the public-What Every Citizen Should Know》	CQ press	1998
James H. Watt & Sjef A. van den Berg	《Research Methods for Communication Science》	Allyn & Bacon	1995
Larry L. Orr	《Socail Experiments -Evaluation Public Programs with Experimental Methods》	SAGE Publications, Inc.	1998
Margaret H. De-Fleur	《Computer-Assisted Investigative Reporting- Development and Methodology》	Lawrence Eribaum Associates, publishers, Mahwah, New Jersey,	1997
McCombs, Maxwell E., Cole, Richard R., Stevenson, Robert L. and Shaw, Donald L.	《Precision Journalism: An Emerging Theory and Technique of News Reporting》	Gazette	1981
Nataliel L. Sproull	《Handbook of Research Methods- A Guide for Practitioners and Students in the Social Sciences》Second Edition	the Scarecrow Press, Inc. Metuchen, N. J., & London,	1995

Nick Moon	《**Opinion polls- History, theory and practice**》	Manchester University Press	1994
Norman M. Bradburn Seymour Sudman	《**Polls & Surveys: Understanding what they tell us**》	Jossey-Bass Publishers	1988
Roger D. Wimmer & Joseph R. Dominick	《**Mass Media Research - An Introduction**》**Fifth Edition**	Wadsworth Publishing Company	1997
Russell A. Jones	《**Research Methods in the Social and Behavioral Sciences**》	Sinauer Associates, Inc.	1996
Sheldon R. Gawiser and G. Evans Witt	《**A Journalist's Guide to Public Opinion Polls**》	Praeger Publishers	1994
W. Lawrence Neuman	《**Social Research Methods-Qualitative and Quantitative Approaches**》**Third Edition**	Allyn & Bacon	1997
北京新闻学会调查组编	《**北京读者、听众、观众调查**》	工人出版社	1985年
范伟达主编	《**世纪图景—21世纪国人生活权威调查**》	中国社会出版社	2000年10月
风笑天	《**现代社会调查方法**》	华中理工大学出版社	1996年10月
风笑天著	《**社会学研究方法**》	中国人民大学出版社	2001年3月
冯士雍 倪加勋 邹国华	《**抽样调查理论与方法**》	中国统计出版社	1998年3月
冯士雍,施锡铨	《**抽样调查—理论、方法与实践**》	上海科学技术出版社	1996年

国家统计局中国经济景气监测中心　中央电视台经济部《中国财经报道》编	《中国都市生活报告》	中国统计出版社	1995年5月
何光先著	《现代新闻学》	重庆出版社	1991年
姜秀珍	《新闻统计学》	新华出版社	1998年5月
金勇进	《非抽样误差分析》	中国统计出版社	1996年2月
柯惠新　丁立宏	《市场调查与分析》	中国统计出版社	2000年
柯惠新　黄京华　沈浩	《调查研究中的统计分析方法》	北京广播学院出版社	1992年
李沛良	《社会研究中的统计应用》	社会科学文献出版社	2001年2月
李文刚著	《用数字说话—数字新闻的理论与操作》	天津人民出版社	2000年11月
刘建鸣　胡运芳主编	《收视率透视》	中国广播电视出版社	2000年11月
刘明华著	《西方新闻采访与写作》	中国人民大学出版社	1993年7月
罗文辉著	《精确新闻报道》	正中书局	1991年4月
倪加勋	《抽样调查》	东北财经大学出版社	1994年11月
苏衡	《传播研究调查法》	三民书局	1986年
陶涵主编	《新闻学、传播学新名词词典》	经济日报出版社	1997年



王波编著	《计算机辅助新闻学概论》	新华出版社	2000年9月
王庆石　卢兴主编	《统计学案例教材》	东北财经大学出版社	1999年12月
王石番	《传播内容分析法—理论与实证》	幼狮文化事业公司	1991年
谢邦昌原著，张尧庭，董麓改编	《抽样调查的理论极其应用方法》	中国统计出版社	1998年
杨孝荣	《社会研究实务》	国立编译馆出版正中书局印行	1995年
袁方　林斌	《社会调查原理与方法》	高等教育出版社	1990年1月
袁亚愚主编	《普通社会学教程》	四川大学出版社	1997年11月
袁卫，吴喜之，贾俊平	《描述统计学》	中国统计出版社	1996年1月
柯惠新，刘红鹰	《民意调查实务》	中国经济出版社	1996年1月
三味工作室	《SPSS10.0 for Windows 实用基础教程》	北京希望电子出版社	2001年2月
苏金明，傅荣华，周建斌，张莲花	《SPSS for Windows 实用指南》	电子工业出版社	2000年9月
郑杭生主编	《社会学概论新修》	中国人民大学出版社	1994年1月

http://www.unc.edu/~pmeyer/meyeres.htm



后　记

经过近一年的努力，《精确新闻学》一书终于完稿了。希望我们的努力能对新闻学的学生以及新闻从业人员有所帮助。

全书共分十二章，第一章介绍了精确新闻学的发展历史以及现状；第二章介绍了精确新闻学中涉及的概念；然后按照精确新闻报道的程序分别介绍了抽样、内容分析方法、实验方法、调查方法、问卷设计、数据的整理和统计分析，最后介绍了精确新闻的写作。考虑到精确新闻报道中最常用的数据收集方法就是电脑辅助的电话调查 CATI，而 CATI 在我国还是新生事物，所以特别设了一章专门来介绍 CATI。另外，对于数据分析来讲，没有统计软件的帮助是无法完成的，所以专设了一章来介绍统计软件 SPSS。肖明策划了本书的写作，并写了第一章到第七章，丁迈写了第八章到第十二章。

曹璐、柯惠新两位博导百忙之中为本书写了序言；香港城市大学祝建华教授、台湾辅仁大学谢邦昌教授提供了自己的研究资料；丁俊杰教授、雷跃杰教授对本书的写作给予了鼓励；书中引用了曹培鑫的硕士毕业论文作为例子；家人和同事对本书的写作给予了大力支持。在此一并表示衷心的感谢。

限于作者的学识水平和经验有限，书中错误及不妥之处在



所难免，恳请广大读者不吝赐教。本书可以新闻学系的教材，为方便使用，我们将把本书做成 Power Point 幻灯片形式，有需要者也可以同我们联系。联系方式为 xiaoming@public.bta.net.cn.

作　者

2001 年 10 月 1 日

[ＧｅｎｅｒａｌＩｎｆｏｒｍａｔｉｏｎ]
书名＝精解新闻学
作者＝
页数＝４４９
ＳＳ号＝１０４９９７８９
出版日期＝

封面页
书名页
版权页
前言
目录
正文